PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-230362
(43)Date of publication of application : 29.08.1995

(51)Int.Cl. G06F 3/06
G11B 20/18
G11B 20/18
G11B 20/18

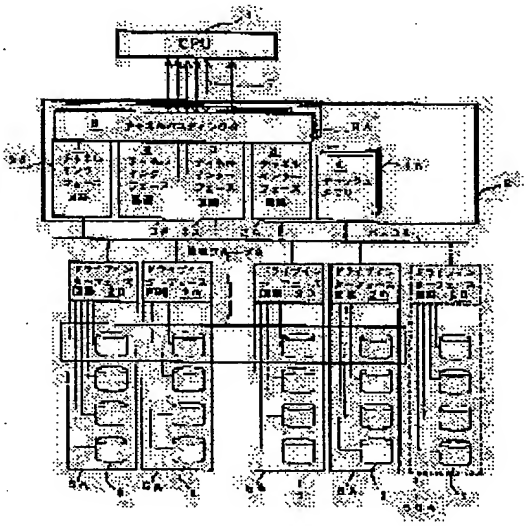
(21)Application number : 06-323920
(22)Date of filing : 30.11.1994

(71)Applicant : HITACHI LTD
(72)Inventor : TSUNODA HITOSHI
TAKAMOTO YOSHIFUMI
KAGIMASA TOYOHIKO

(30)Priority
Priority number : 05329810 Priority date : 30.11.1993 Priority country : JP

(54) DISK ARRAY DEVICE

(57)Abstract:
PURPOSE: To access data in a faulty drive by providing a circuit which writes data, which should be written in the faulty drive,-in an alternative drive and writes data in normal drives on a board at the time when data write to the alternative drive is possible after substitution with the alternative drive of the faulty drive.
CONSTITUTION: If a fault occurs in any drive mounted on a drive board 5A, the faulty drive is not separated from the drive board 5A left connected to a mother board, but the drive board 5A is separated from the mother board with normal drives mounted. After the faulty drive is substituted with a normal alternative drive, the drive board 5A is connected to the mother board again. A logical group 9 is the fault recovery unit, and drives in the logical group 9 hold an error correction data group which consists or m-1 data ana generated parity data.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

(19)日本国特許庁（ＪＰ）　(12)公開特許公報（Ａ）　(11)特許出願公開番号
特開平7－230362
(43)公開日　平成７年(1995)８月29日

(51)Int.Cl.[6]	識別記号	庁内整理番号	ＦＩ	技術表示箇所
Ｇ０６Ｆ　3/06	５４０			
Ｇ１１Ｂ　20/18	５３２　Ｂ	9074－5D		
	５５２　Ａ	9074－5D		
	５７０　Ｚ	9074－5D		

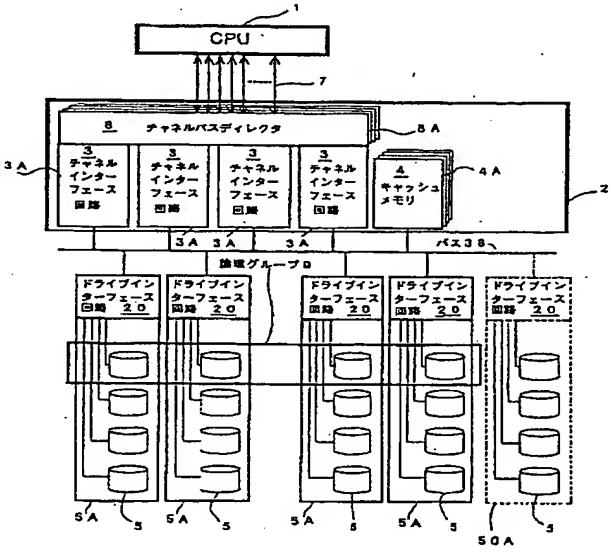
審査請求　未請求　請求項の数35　ＦＤ　（全 25 頁）

(21)出願番号　特願平6－323920
(22)出願日　平成６年(1994)11月30日
(31)優先権主張番号　特願平5－329810
(32)優先日　平５(1993)11月30日
(33)優先権主張国　日本（ＪＰ）

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地
(72)発明者　角田　仁
東京都国分寺市東恋ケ窪一丁目280番地
株式会社日立製作所中央研究所内
(72)発明者　高本　良史
東京都国分寺市東恋ケ窪一丁目280番地
株式会社日立製作所中央研究所内
(72)発明者　鍵政　豊彦
東京都国分寺市東恋ケ窪一丁目280番地
株式会社日立製作所中央研究所内
(74)代理人　弁理士　矢島　保夫

(54)【発明の名称】　ディスクアレイ装置

(57)【要約】　（修正有）
【目的】複数のドライブを基板に搭載し、これを、共通のマザーボードに並置でき、かつ何れかのドライブに障害が生じたときも、保持されたデータを上位装置からアクセス可能にするディスクアレイ装置を提供する。
【構成】アレイコントローラ２は、複数の基板の内、互いに異なるものに搭載されている複数のドライブに書き込む回路と、いずれかのドライブに障害が発生したためにマザーボードからその一つの基板が分離された状態において、その分離されたドライブに保持されているデータが読み出しのために上位装置からアクセスされたときに、他の複数の基板に保持されている複数のドライブから、その一つのデータと同じ誤り訂正データグループに属するデータおよびパリティデータを読み出す回路と、読み出された他の複数のデータと読み出されたパリティデータとから読み出すべきデータを回復する回路とを有する。



【特許請求の範囲】

【請求項１】（ａ）一群のディスク記憶装置であって、該一群のディスク記憶装置は複数の論理グループに分割され、各論理グループは、複数のデータと、それらから生成された少なくとも一つの誤り訂正符号とをそれぞれ有する複数群の誤り訂正データを保持する複数のディスク記憶装置からなるものと、

（ｂ）それぞれ該一群のディスク記憶装置の内の一部の複数のディスク記憶装置を保持する複数の個別基板と、

（ｃ）該一群の複数のディスク記憶装置の各々に分離可能に接続され、上位装置からのデータ読み出し要求またはデータ書き込み要求を実行するアレイコントローラであって、（ｃ１）各データ書き込み要求を、（ｃ１１）そのデータ書き込み要求に付随する書き込みデータが属する誤り訂正データ群に属すべき誤り訂正符号を生成し、（ｃ１２）上記書き込みデータと該生成された誤り訂正符号とを、一つの誤り訂正データ群に属するデータとして、該複数の論理グループの一つに属する複数のディスク記憶装置に書き込むように実行し、（ｃ２）各データ読み出し要求を、（ｃ２１）それが要求するデータを保持する、該一群のディスク記憶装置の一つが正常なときには、そのディスク記憶装置からその要求されたデータを読み出し、（ｃ２２）そのディスク装置に障害があるときには、その障害があるディスク記憶装置が属する論理グループに属する複数のディスク記憶装置の内、上記障害があるディスク装置以外のディスク記憶装置に保持され、該読み出し要求で要求されたデータが属する、誤り訂正データ群に属し、一つの誤り訂正符号と他の複数のデータとから、該要求されたデータを回復するものとを有し、

（ｄ）該一群のディスク記憶装置の内、同一の論理グループに属する複数のディスク記憶装置が、該複数の個別基板の内の互いに異なるものに分散して搭載されているディスクアレイ装置。

【請求項２】該複数の個別基板を着脱可能に保持する、上記複数の個別の基板に共通な基板をさらに有し、

上記アレイコントローラは、該一群の複数のディスク記憶装置に該共通の基板の上の信号線路を介して接続され、

該複数の個別基板は、該共通の基板にほぼ垂直に、かつ、互いにほぼ平行に保持されている請求項１記載のディスクアレイ装置。

【請求項３】各個別基板に搭載されている複数のディスク記憶装置は、互いに異なる複数の論理グループに属する請求項１記載のディスクアレイ装置。

【請求項４】各個別基板に搭載されている複数のディスク記憶装置は、互いに異なる複数の論理グループに属する請求項２記載のディスクアレイ装置。

【請求項５】上記アレイコントローラは、

該障害がある一つのディスク記憶装置を搭載した上記複数の個別基板の一つが、該共通の基板から分離されている後に上記上位装置から供給され、かつ、その障害があるディスク記憶装置に保持された一つのデータを要求する一つの読み出し要求を受理し、

その受理した読み出し要求の実行にあっては、その障害があるディスク記憶装置が属する論理グループに属し、その一つの基板以外の他の複数の基板に保持された他の複数のディスク記憶装置に保持された他の複数のデータからその要求されたデータを回復し、回復されたデータを上記上位装置に供給し、

さらに、上記分離の後に該上位装置から供給され、かつ、該一つの基板に保持された他の正常なディスク記憶装置に保持されたデータを要求する他の読み出し要求を受理し、

その受理した他の読み出し要求の実行にあっては、その正常な他のディスク記憶装置が属する他の論理グループに属し、その一つの基板以外の他の複数の基板に保持されたさらに他の複数のディスク記憶装置に保持されたさらに他の複数のデータからその要求された他のデータを回復し、回復された他のデータを上記上位装置に供給する請求項２記載のディスクアレイ装置。

【請求項６】ランダムアクセス可能なメモリをさらに有し、

該アレイコントローラは、

該分離の後に上記上位装置から供給され、該分離された一つの個別基板に保持されている複数の正常なディスク記憶装置もしくは該障害があるディスク記憶装置のいずれか一つへデータを書き込むことを要求する複数の書き込みを受理し、

該書き込みデータを該メモリに書き込み、

該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換された後、該メモリに保持された複数の書き込みデータを用いて該受理された複数の書き込み要求を実行する請求項５記載のディスクアレイ装置。

【請求項７】上記アレイコントローラは、該一つの個別基板が該共通に基板に再度接続された後に、上記ランダムアクセスメモリに保持された複数の書き込みデータの書き込みの前に、該一つの個別基板以外の該他の複数の基板に搭載された他の複数のディスク記憶装置内の複数のデータおよび複数の誤り訂正符号から、該障害があるディスク記憶装置に保持されていた複数のデータを回復し、該回復された複数のデータを該交替のディスク記憶装置に書き込む請求項６記載のディスクアレイ装置。

【請求項８】該アレイコントローラは、

該一つの個別基板が該共通の基板から分離されている状態において、上記上位装置から供給された上記書き込み要求を受理し、

その一つの個別基板が再度その共通の基板に接続されるを待たないで、その受理した書き込み要求を実行し、

その実行にあたっては、その書き込みデータに対して誤

り訂正符号を生成し、その障害があるディスク記憶装置もしくはその正常なディスク記憶装置の一つと同じ論理グループに属し、その一つの個別基板以外の他の一つの個別基板に保持された一つのディスク記憶装置に生成した誤り訂正符号を書き込み、
該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換され、該分離された一つの個別基板が、該共通の基板に再度接続されたときに、該一つの個別基板に搭載された該交替のディスク記憶装置および該複数の正常なディスク記憶装置の各々に書き込まれるべき複数のデータを、その一つの個別基板以外の他の複数の個別基板に保持された他の複数のディスク記憶装置に保持された他の複数のデータから回復し、
回復された複数のデータを、その、各ディスク記憶装置に書き込む請求項５記載のディスクアレイ装置。
【請求項９】該アレイコントローラは、
該一つの個別基板に保持された複数の正常なディスク記憶装置の各々に関しては、上記一つの個別基板が該共通基板から分離された後に上記上位装置から供給された複数の書き込み要求が指定した複数の書き込みデータを回復し、
該一つの個別の基板が該共通の基板から分離される前に該各正常なディスク記憶装置にすでに保持されていた書き込みデータは回復しない請求項８記載のディスクアレイ装置。
【請求項１０】該アレイコントローラに接続された少なくとも一つの予備のディスク記憶装置をさらに有し、
該アレイコントローラは、
該一つの個別基板が該共通の基板から分離された状態で、該上位装置から供給された、該分離された一つの個別基板に搭載された該障害があるディスク記憶装置もしくは複数の正常なディスク記憶装置に書き込むべきデータを、上記予備のディスク記憶装置に書き込み、
該上位装置からその後供給された読み出し要求が指定するデータが該予備のディスク記憶装置にすでに書き込まれているとき、該予備のディスク記憶装置からその要求されたデータを読み出し、該上位装置に供給する請求項５のディスクアレイ装置。
【請求項１１】該アレイコントローラは、
該一つの個別基板が該共通の基板から分離された状態で、該上位装置から供給された読み出し要求を実行したときに、その読み出し要求が、該分離された一つの個別基板に搭載された、該障害があるディスク記憶装置もしくは複数の正常なディスク記憶装置に保持されたデータの読み出しを要求するとき、その読み出し要求の実行時に回復したデータを該予備のディスク記憶装置に書き込み、
該読み出し要求が指定するデータの読み出し要求がその後該上位装置から要求されたときに、該予備のディスク記憶装置からその要求されたデータを読み出す請求項１０記載のディスクアレイ装置。
【請求項１２】該アレイコントローラは、該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換された後、該分離された一つの個別基板が、該共通の基板に再度接続されたときに、上記予備ディスク記憶装置に保持された複数の書き込みデータの内、その個別基板が該共通基板から分離されていた間に該上位装置から供給された、その個別の基板に搭載された複数の正常なディスク装置に書き込まれるべきであった複数のデータを、該複数の正常なディスク記憶装置に転送する請求項１０記載のディスクアレイ装置。
【請求項１３】該アレイコントローラは、該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換された後、該分離された一つの個別基板が、該共通の基板に再度接続されたときに、該障害ドライブに障害が発生する前に該障害があるディスク記憶装置にすでに保持されていたデータおよび該障害があるディスク記憶装置に障害が発生した後に該上位装置から供給された、その障害があるディスク記憶装置に書き込むべきであった他の複数の書き込みデータとを、該一つの個別基板以外の複数の個別基板に搭載された複数のディスク記憶装置に記憶された複数のデータと複数の誤り訂正符号から回復し、回復された複数のデータを該交替のディスク記憶装置に書き込む請求項１２記載のディスクアレイ装置。
【請求項１４】該アレイコントローラに接続された複数の予備のディスク記憶装置をさらに有し、
該アレイコントローラは、
該一つの個別基板が該共通の基板から分離されたときに、該一つの個別基板に保持されている複数のディスク記憶装置の各々に対応して、そのディスク記憶装置の代わりに使用するディスク記憶装置を該複数の予備のディスク記憶装置から選択し、
該一つの個別基板が該共通の基板から分離された状態で、該上位装置から供給された、該分離された一つの個別基板に搭載された該複数のディスク記憶装置のいずれか一つに書き込みデータを書き込むことを要求する書き込み要求を受理し、
その受理した書き込み要求の実行にあっては、上記複数の予備のディスク記憶装置の内の、その一つのディスク記憶装置記憶に対応して選択されたディスク記憶装置にその書き込みデータを書き込み、
該上位装置からその後供給された読み出し要求が指定するデータが上記複数の予備のディスク記憶装置のいずれかに保持されているとき、その予備のディスク記憶装置からその要求されたデータを読み出し、該上位装置に供給する請求項５記載のディスクアレイ装置。
【請求項１５】該共通の基板に着脱自在に、かつ、該複数の個別の基板に略平行に保持された少なくとも一つの予備の個別の基板をさらに有し、

該複数の予備のディスク記憶装置は、該予備の基板に搭載され、該共通の基板の上記信号線路を介して該アレイコントローラに接続されている請求項１４記載のディスクアレイ装置。

【請求項１６】該アレイコントローラは、

該一つの個別基板が該共通の基板から分離された状態で、該上位装置から供給された読み出し要求を実行したときに、その読み出し要求が、該分離された一つの個別基板に搭載された、該障害があるディスク記憶装置もしくは複数の正常なディスク記憶装置に保持されたデータの読み出しを要求するとき、その読み出し要求の実行により回復したデータを該予備のディスク記憶装置に書き込み、

該読み出し要求が指定するデータの読み出し要求がその後該上位装置から要求されたときに、該予備のディスク記憶装置からその要求されたデータを読み出す請求項１４記載のディスクアレイ装置。

【請求項１７】該アレイコントローラは、該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換され、該分離された一つの個別基板が、該共通の基板に再度接続されたときに、上記複数の予備ディスク記憶装置の内、該一つの個別基板に搭載された複数の正常なディスク記憶装置に対応して選択された複数の予備のディスク記憶装置に保持された複数の書き込みデータを、該複数の正常なディスク記憶装置に書き込む請求項１４記載のディスクアレイ装置。

【請求項１８】該アレイコントローラは、

該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換され、該分離された一つの個別基板が、該共通の基板に再度接続されたときに、該一つの個別基板が該共通基板から分離される前に該上位装置から供給された該障害があるディスク記憶装置に書き込むべきであった複数の書き込みデータおよび該一つの個別基板が該共通基板から分離される前に該上位装置から供給された該障害があるディスク記憶装置に書き込むべきであった他の複数の書き込みデータとを、該一つの個別基板以外の複数の個別基板に搭載された複数のディスク記憶装置に保持された、複数のデータおよび複数の誤り訂正符号から回復し、

回復された複数のデータを該交替のディスク記憶装置に書き込む請求項１４記載のディスクアレイ装置。

【請求項１９】該アレイコントローラは、

該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換され、該分離された一つの個別基板が、該共通の基板に再度接続されたときに、該障害のあるドライブが障害状態になる前に該障害があるドライブが保持していた複数のデータを回復し、

回復された複数のデータを該交替のディスク記憶装置に書き込み、

その書き込み該障害があるディスク記憶装置に対して選択された該一つの予備のディスク記憶装置に保持されている、該一つの個別基板が該共通基板から分離された後に該上位装置から供給された該障害があるディスク記憶装置に書き込むべきであった複数の書き込みデータを、該交替ディスク記憶装置に転送する請求項１４記載のディスクアレイ装置。

【請求項２０】該アレイコントローラは、

該障害のあるディスク記憶装置が現に障害になった後に、該一つの個別基板上の該障害のあるディスク記憶装置が現に障害になる前に、その障害ドライブが保持していた複数のデータを、該一つの個別基板以外の複数の個別基板に搭載された複数のディスク記憶装置に保持された複数のデータから回復し、

回復された複数のデータを該障害のあるディスク記憶装置に対応して選択された予備のディスク記憶装置に書き込み、

該回復されたデータの書き込みが完了後、該上位装置から供給される、該障害があるディスク記憶装置に対する書き込み要求および読み出し要求を、該一つの予備のディスク記憶装置に対して実行するように、該読み出し要求および該書き込み要求の実行する請求項１７記載のディスクアレイ装置。

【請求項２１】該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換された後、該分離された一つの個別基板が、該共通の基板に再度接続されたときに、該障害があるディスク記憶装置に対応して選択された該一つの予備のディスク記憶装置を、該障害があるディスク記憶装置に代わりのディスク記憶装置として使用し、

該交替のディスク記憶装置を新たな予備のディスク記憶装置として使用する請求項２０記載のディスクアレイ装置。

【請求項２２】該共通の基板に着脱自在に、かつ、該複数の個別の基板に略平行に保持された少なくとも一つの予備の個別の基板をさらに有し、

該複数の予備のディスク記憶装置は、該予備の基板に搭載され、該共通の基板の上記信号線路を介して該アレイコントローラに接続されている請求項２１記載のディスクアレイ装置。

【請求項２３】該アレイコントローラは、

該一つの個別基板が該共通の基板から分離されたときに、該一つの個別基板上の該障害のあるディスク記憶装置が保持していた書き込みデータを、該一つの個別基板以外の複数の個別基板に搭載された複数のディスク記憶装置に保持された複数のデータおよび複数の誤り訂正符号から回復し、

回復された複数のデータを該障害のあるディスク記憶装置に対応して選択された予備のディスク記憶装置に書き込み、

該回復されたデータの書き込みが完了後、該上位装置か

ら供給される、該障害があるディスク記憶装置に保持されているデータに対する書き込み要求および読み出し要求を、該一つの予備のディスク記憶装置に対して実行するように、該読み出し要求および該書き込み要求の実行する請求項１７記載のディスクアレイ装置。

【請求項２４】該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換された後、該分離された一つの個別基板が、該共通の基板に再度接続されたときに、該障害があるディスク記憶装置に対応して選択された該一つの予備のディスク記憶装置を、該障害があるディスク記憶装置に代わりのディスク記憶装置として使用し、

該交替のディスク記憶装置を新たな予備のディスク記憶装置として使用する請求項２０記載のディスクアレイ装置。

【請求項２５】該共通の基板に着脱自在に、かつ、該複数の個別の基板に略平行に保持された少なくとも一つの予備の個別の基板をさらに有し、

該複数の予備のディスク記憶装置は、該予備の基板に搭載され、該共通の基板の上記信号線路を介して該アレイコントローラに接続されている請求項２４記載のディスクアレイ装置。

【請求項２６】該アレイコントローラは、

該上位装置から供給された書き込み要求に付随する一つの書き込みデータを複数のサブデータに分割し、該生成された複数のサブデータに対する誤り訂正符号を生成し、該複数のサブデータと該生成された誤り訂正符号を一つの誤り訂正データ群に属するデータとして、互いに異なる個別基板に搭載された複数のディスク記憶装置に書き込むように、該書き込み要求を実行し、

さらに、該上位装置から供給されたデータ読み出し要求が指定するデータを構成する複数のサブデータを、異なる個別基板に搭載された複数のディスク記憶装置から読み出し、読み出された複数のサブデータを結合し、結合後のデータを該上位装置に供給するように、各読み出し要求を実行する請求項４記載のディスクアレイ装置。

【請求項２７】各誤り訂正データ群は、該上位装置から供給された複数の書き込みデータと、該複数の書き込みデータに対する誤り訂正符号とからなり、

該アレイコントローラは、

上位装置から供給された各書き込み要求に実行にあっては、その書き込み要求に付随する書き込みデータが属する誤り訂正符号データ群に対してすでに決定された旧の誤り訂正符号と、該書き込みデータで更新される旧書き込みデータとを、それぞれ互いに異なる個別基板に保持された相異なるディスク記憶装置から読み出し、

読み出された旧誤り訂正符号と読み出された旧書き込みデータと該書き込みデータとから、該書き込みデータでもって該旧書き込みデータを更新した後の誤り訂正符号を生成し、

その書き込みデータと、その生成された誤り訂正符号を互いに異なる個別基板に搭載された複数のディスク記憶装置に書き込み、

上位装置から供給された読み出し要求の実行にあっては、その読み出し要求が指定するデータを、いずれかひとつの個別基板に搭載された一つのディスク記憶装置から読み出し、読み出されたデータを該上位装置に供給する請求項４記載のディスクアレイ装置。

【請求項２８】上記アレイコントローラは、

該障害がある一つのディスク記憶装置を搭載した上記複数の個別基板の一つが、該共通の基板から分離されている後に上記上位装置から供給され、かつ、その障害があるディスク記憶装置に保持された一つのデータを要求する一つの読み出し要求を受理し、

その受理した読み出し要求の実行にあっては、その障害があるディスク記憶装置が属する論理グループに属し、その一つの基板以外の他の複数の基板に保持された他の複数のディスク記憶装置に保持された他の複数のデータからその要求されたデータを回復し、回復されたデータを上記上位装置に供給し、

さらに、上記分離の後に該上位装置から供給され、かつ、該一つの基板に保持された他の正常なディスク記憶装置に保持されたデータを要求する他の読み出し要求を受理し、

その受理した他の読み出し要求の実行にあっては、その正常な他のディスク記憶装置が属する他の論理グループに属し、その一つの基板以外の他の複数の基板に保持されたさらに他の複数のディスク記憶装置に保持されたさらに他の複数のデータからその要求された他のデータを回復し、回復された他のデータを上記上位装置に供給する請求項１記載のディスクアレイ装置。

【請求項２９】該アレイコントローラは、

該分離の後に上記上位装置から供給され、該分離された一つの個別基板に保持されている複数の正常なディスク記憶装置もしくは該障害があるディスク記憶装置のいずれか一つへデータを書き込むことを要求する複数の書き込み要求を受理し、

該書き込みデータを該メモリに書き込み、

該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換された後、該メモリに保持された複数の書き込みデータを用いて該受理された複数の書き込み要求を実行する請求項２８記載のディスクアレイ装置。

【請求項３０】該アレイコントローラに接続された複数の予備のディスク記憶装置をさらに有し、

上記アレイコントローラは、

該一つの個別基板が該共通の基板から分離されたときに、該一つの個別基板に保持されている複数のディスク記憶装置の各々に対応して、そのディスク記憶装置の代わりに使用するディスク記憶装置を該複数の予備のディ

スク記憶装置から選択し、
該一つの個別基板が該共通の基板から分離された状態で、該上位装置から供給された、該分離された一つの個別基板に搭載された該複数のディスク記憶装置のいずれか一つに書き込みデータを書き込むことを要求する書き込み要求を受理し、
その受理した書き込み要求の実行にあっては、上記複数の予備のディスク記憶装置の内の、その一つのディスク記憶装置記憶に対応して選択されたディスク記憶装置にその書き込みデータを書き込み、
該上位装置からその後供給された読み出し要求が指定するデータが上記複数の予備のディスク記憶装置のいずれかに保持されているとき、その予備のディスク記憶装置からその要求されたデータを読み出し、該上位装置に供給する請求項２８記載のディスクアレイ装置。
【請求項３１】一群のディスク記憶装置であって、該一群のディスク記憶装置は複数の論理グループに分割され、各論理グループは、複数のデータと、それらから生成された少なくとも一つの誤り訂正符号とをそれぞれ有する複数群の誤り訂正データを保持する複数のディスク記憶装置からなるものと、
それぞれ該一群のディスク記憶装置の内の一部の複数のディスク記憶装置を保持する複数の個別基板と、
該一群の複数のディスク記憶装置の各々に分離可能に接続され、上位装置からのデータ読み出し要求またはデータ書き込み要求を実行するアレイコントローラとを有するディスクアレイ装置において、
（ａ）該一群のディスク記憶装置の内、同一の論理グループに属する複数のディスク記憶装置は、該複数の個別基板の内の互いに異なるものに分散して搭載された複数のディスク記憶装置から構成されるように、各論理グループに属するディスク記憶装置を決定し、
（ｂ）いずれかのディスク記憶装置に障害が発生したとき、そのディスク記憶装置を含む一つの個別基板を該アレイコントローラから分離し、
（ｃ）該アレイコントローラから分離されている状態で上記上位装置にから発行され、該分離された一つの個別の基板上の上記障害が発生したディスク記憶装置および他の正常なディスク記憶装置の内のいずれか一つに保持されたデータを要求する複数の読み出し要求を受理し、
（ｄ）各受理した読み出し要求が要求する読み出しデータを、その一つの個別基板以外の基板に保持された複数のディスク記憶装置から回復し、回復されたデータを該上位装置に供給するように、各読み出し要求を実行するディスクアレイ装置のアクセス方法。
【請求項３２】該アレイコントローラはランダムアクセス可能なメモリをさらに有し、　　該方法は、
該アレイコントローラから分離されている状態で上記上位装置にから発行され、該分離された一つの個別基板に保持されている該障害があるディスク記憶装置もしくは複数の正常なディスク記憶装置のいずれか一つへ書き込むべきデータを指定する書き込み要求を受理し、
該受理した書き込み要求が指定する書き込みデータを、該メモリに一時的に保持し、
該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換された後に、該一つの個別基板が該アレイコントローラに再度接続されたときに、該メモリに保持された該書き込みデータを用いて該一つの書き込み要求を実行するステップをさらに有する請求項３１記載のディスクアレイ装置のアクセス方法。
【請求項３３】上記アレイディスク装置は、該アレイコントローラに接続された少なくとも一つの予備のディスク記憶装置をさらに有し、
上記方法は、
該一つの個別基板が該共通の基板から分離された状態で、該上位装置から受理され、該分離された一つの個別基板に搭載された該障害があるディスク記憶装置もしくは他の正常なディスク記憶装置のいずれか一つに書き込むべきデータを指定する複数の書き込み要求を受理し、
該書き込み要求の各々を、それが指定する書き込みデータを上記予備のディスク記憶装置に書き込むように実行し、
該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換され、該一つの個別基板が該アレイコントローラに再度接続された後に、該予備のディスク記憶装置に保持された書き込みデータの内、上記一つの個別基板に保持された該他の正常なディスク装置に書き込まれるべきデータを、該他の正常なディスク記憶装置に転送するステップをさらに有する請求項３１記載のディスクアレイ装置のアクセス方法。
【請求項３４】上記アレイディスク装置は、該アレイコントローラに接続された複数の予備のディスク記憶装置をさらに有し、
上記方法は、
該一つの個別基板の上の各ディスク装置に対応して、該複数の予備のディスク記憶装置の一つを選択し、
該一つの個別基板が該アレイコントローラから分離されるときに、該障害が発生した一つのディスク記憶装置が保持していたデータを、該一つの個別基板以外の他の個別基板に保持された複数のデータおよび複数の誤り訂正符号から回復し、
回復された複数のデータを該一つの障害が発生したディスク記憶装置に対応して選択された、一つの予備のディスク記憶装置に転送し、
該一つの個別基板が該アレイコントローラから分離された状態で、該上位装置から受理したいずれかの書き込み要求が、該分離された一つの個別基板に搭載された該障害があるディスク記憶装置もしくは複数の正常なディスク記憶装置のいずれか一つに書き込むべきデータを指定するとき、上記複数の予備のディスク記憶装置の内、そ

の一つのディスク記憶装置に対応して選択された一つの予備のディスク記憶装置に書き込み、
該障害があるディスク記憶装置が正常な交替のディスク記憶装置により置換された後に、該一つの個別基板が該アレイコントローラに再度接続されたときに、該一つの予備のディスク記憶装置を、該障害が発生した一つのディスク記憶装置の代わりに使用し、
該交替のディスク記憶装置を新たな予備のディスク記憶装置として使用するステップをさらに有する請求項３１記載のディスクアレイ装置のアクセス方法。
【請求項３５】該一つの個別基板が該アレイコントローラに再度接続された後に、該複数の予備のディスク記憶装置の内、該他の正常なディスク記憶装置に対して選択された他の一つの予備のディスク記憶装置に保持された書き込みデータを、該他の正常なディスク記憶装置に転送し、
上記他の予備のディスク記憶装置を新たに予備のディスク記憶装置として解放するステップをさらに有する請求項３４記載のディスクアレイ装置のアクセス方法。
【発明の詳細な説明】
【０００１】
【産業上の利用分野】本発明は、複数のデータを誤り訂正符号とともに保持する複数のディスク装置を有するディスクアレイ装置に関する。
【０００２】
【従来の技術】現在のコンピュータシステムにおいては、ＣＰＵ等の上位側が必要とするデータは２次記憶装置に格納され、ＣＰＵが必要とするときに２次記憶装置に対してデータの書き込み、読み出しを行うようになっている。この２次記憶装置としては、一般に不揮発な記憶媒体が使用され、代表的なものとして磁気ディスク装置や光ディスクなどがあげられる。以下、これらに例示されるディスク装置をドライブと呼ぶ。
【０００３】近年の高度情報化に伴い、コンピュータシステムにおいて、２次記憶装置の高性能化が要求されてきた。その一つの解として、多数の比較的容量の小さなドライブにより構成されるディスクアレイが考えられている。
【０００４】例えば、D. Patterson, G. Gibson, and R. H. Kartz;A Case for Redundant Arraysof Inexpensive Disks(RAID), ACM SIGMD Conference, Chicago, IL, (June 1988), pp. 109-116（以下、第１の従来技術と呼ぶ）なる論文においては、データを分割して並列に処理を行うディスクアレイ（レベル３）とデータを分散して独立に扱うディスクアレイ（レベル４，５）について、それらの性能および信頼性の検討結果が報告されている。
【０００５】レベル３のディスクアレイは、ＣＰＵから一つの書き込み要求に付随して転送されてきた一つの書き込みデータを複数に分割し、分割後の複数のデータから誤り訂正用のパリティデータを作成する。そして、これらのデータ（分割後の複数のデータおよびパリティデータ）を複数のドライブに並列に格納する。また、ＣＰＵからデータの読み出しが指示された場合は、まず各々のドライブから分割されたデータを並列に読み込み、それらを結合してＣＰＵへ転送する。
【０００６】レベル４，５のディスクアレイは、ＣＰＵから転送されてきた複数の書き込み要求に付随する複数のデータを上記分割したデータの代わりに使用すること、およびデータの読み出しは、個々のデータに対して行なわれ、レベル３でのデータの結合は行なわれない点で、レベル３の場合と異なる。
【０００７】複数のデータと、それらから生成された誤り訂正用のデータとは、「誤り訂正データ群」あるいは「パリティグループ」あるいは「パリティデータグループ」と呼ばれることがある。誤り訂正用のデータとしてはパリティデータのほかにも各種の誤り訂正用データを用いることができるが、本明細書でも、「パリティグループ」との用語を誤り訂正用データがパリティデータでない場合にも用いるものとする。また、以下では誤り訂正符号としては、パリティデータの場合について説明するが、本願発明はこのパリティデータ以外にも適用可能である。
【０００８】以下では、同一のパリティグループのデータを格納する複数のドライブを「論理グループ」と呼ぶ。論理グループは障害回復の単位であり、論理グループ内の何れかのドライブに障害が発生した場合はその論理グループ内の他のドライブのデータから障害回復を行うことができる。
【０００９】例えば、分割後のデータを格納したドライブの中の１台に障害が発生し、データが読み出せなくなった場合は、残りのドライブ内のデータとパリティデータとから、障害が発生したドライブ内のデータを復元することができる。ディスクアレイのような多数のドライブにより構成される装置では、部品点数が増加することにより、障害が発生する確率が高くなる。そこで、ディスク装置としての信頼性の向上を図る目的で、このようにパリティを用意して障害の回復を行うことができるようにしている。
【００１０】また、何れかの障害ドライブを他の交替ドライブと置換するまでの間、その障害ドライブに書き込むべきデータを、あらかじめ設けられた予備のドライブに書込むことも提案されている。例えば、米国特許５，０７７，７３６（以下、第２の従来技術と呼ぶことがある）に開示されたものである。
【００１１】従来の、典型的なレベル３のディスクアレイ装置（以下、第３の従来技術と呼ぶことがある）では、複数のドライブは、ケーブルでもってアレイコントローラに接続される。その際、異なる論理グループの各々を構成するドライブの数に等しい数のバスでもってそれらのドライブをアレイコントローラに接続していた。

すなわち、異なる論理グループの第１のドライブが、共通の第１のバスにてアレイコントローラに接続されるように、それらのドライブはケーブルでもってデージチェイン状に接続される。同様に、それらの論理グループの第２のドライブが、共通の第２のバスにてアレイコントローラに接続されるように、それらのドライブはケーブルでもってデージチェイン状に接続される。他のドライブも同様である。このようなケーブルを使用したドライブの接続は、装置を構成するドライブの総数が多くなるときには、装置専有面積あるいは保守の容易さなどの点であまり望ましくない。したがって、小型のドライブを多数実装する方法が望まれる。

【００１２】特開平３－１０８１７８号（米国特許出願番号第４０９４９５号（１９８９年９月１９日出願）に対応する）（以下、第４の従来技術と呼ぶことがある）には、底面に多数のピンを有する複数のドライブを半導体基板に搭載する技術が開示されている。この技術は、小型のドライブを多数基板上に実装することを目的としている。しかし、この従来技術には、これらのドライブにより論理グループを構成するか否か、構成するときにはどのようなドライブにより各論理グループを構成するかについては開示されていない。

【００１３】特開平４－２２８１５３号（米国特許出願番号第５０２２１５号（１９９０年３月３０日出願）に対応する）（以下、第５の従来技術と呼ぶことがある）には、それぞれ複数のドライブを共通の配線付きの基板の一つの面に着脱可能に搭載した複数の装置（アレイと呼ばれている）を使用したディスクアレイが開示されている。各アレイ内の複数のドライブは、同一の論理グループに属する。何れかの基板の何れかのドライブに障害が生じたとき、そのドライブを搭載した基板をシステムに接続したままに保持し、その障害が生じたドライブを他の交替ドライブと置換するためにその基板から分離する。しかし、この状態でも、その障害ドライブへのアクセスが可能になっている。すなわち、その障害ドライブに保持されているデータの読み出し要求が発生したとき、その論理グループ内の他のドライブ内のデータとパリティデータとから、そのアクセス対象となったデータを回復して読み出す。

【００１４】

【発明が解決しようとする課題】上記第５の従来技術では、複数のドライブを共通の基板に搭載したアレイを共通のマザーボードに搭載するので、多くの小型のドライブを高密度に実装することが可能になる。しかし、何れかのドライブに障害が発生したときに、その障害ドライブを他の交替ドライブと置換するために、その障害ドライブがそれを搭載した基板から分離できなければならない。この障害ドライブの分離のためには、各ドライブの上方に、ユーザが手でそのドライブをアクセスできる空間を残す必要がある。したがって、この空間のために、このような基板を狭い間隔で多数配列するには限度がある。

【００１５】したがって、本発明の目的は、複数のドライブを複数の基板に搭載し、このような複数の基板を、共通のマザーボードに比較的狭い間隔で並置でき、それでいて何れかのドライブに障害が生じたときでも、その障害ドライブに保持されたデータを上位装置からアクセス可能にするディスクアレイ装置を提供することにある。

【００１６】

【課題を解決するための手段】この目的達成のために、本発明によるディスクアレイ装置は、マザーボードと、該マザーボードの一つの面に着脱可能に保持され、互いにほぼ平行に配置された複数の基板と、各基板の一つの面に搭載された複数のドライブと、該マザーボード上に設けられた信号線路を介して該複数の基板に搭載された複数のドライブからのデータの読み出しおよびそれらへのデータの書き込みを制御するアレイコントローラとを有し、該アレイコントローラは、同一の誤り訂正データグループに属する、複数のデータおよびそれらの複数のデータに対して使用する誤り訂正コードを、上記複数の基板の内、互いに異なるものに搭載されている複数のドライブに書き込む回路と、上記基板のいずれか一つの上に搭載されたいずれかのドライブに障害が発生したために該マザーボードからその一つの基板が分離された状態において、その分離された一つの基板の上のドライブに保持されているデータが読み出しのために上位装置からアクセスされたときに、そのドライブが障害ドライブの場合でもあるいは正常なドライブでも、それぞれその一つの基板以外の他の複数の基板に保持されている複数のドライブから、その一つのデータと同じ誤り訂正データグループに属する他の複数のデータおよびパリティデータを読み出す回路と、読み出された他の複数のデータと読み出されたパリティデータとから該読み出すべき一つのデータを回復する回路とを有する。

【００１７】本願発明の第１の望ましい態様では、上記障害ドライブを搭載した一つの基板がマザーボードから分離されている状態で、その一つの基板に搭載されたいずれか一つのドライブに書き込むべきデータが上位装置から供給されたときに、そのデータを一時的に保持するランダムアクセス可能なメモリと、上記障害ドライブがいずれかの交替ドライブにより置換された後、上記一つの基板が上記マザーボードに再度接続され、その交替ドライブがデータ書き込み可能になったときに、上記メモリに保持された書き込みデータの内、上記障害ドライブに書き込むべきデータを該交替ドライブに書き込み、該分離されていた基板上の正常なドライブに書き込むべきデータを、その正常なドライブに書き込む回路を有する。

【００１８】本発明の他の望ましい態様では、予め上記

データを保持するためのドライブが搭載されている基板と異なる他の基板上に設けられた、少なくとも一つの基板上の搭載されるデータ保持用のドライブの数以上の予備のドライブと、上記障害ドライブを搭載した一つの基板がマザーボードから分離されたとき、その一つの基板上の各ドライブに対応して、該複数の予備のドライブの一つを選択し、その一つの基板が分離された状態で、その一つの基板に搭載されたいずれか一つのドライブに書き込むべきデータが上位装置から供給されたときに、上記複数の予備のドライブの内、そのドライブに対して選択された予備のドライブにその書き込みデータを書き込む回路と、上記障害ドライブがいずれかの交替ドライブにより置換された後、上記一つの基板が上記マザーボードに再度接続され、その交替ドライブがデータ書き込み可能になったときに、該一つの基板の上の各ドライブに上記複数の予備のドライブの内、その各ドライブに対して選択された予備のドライブに保持された書き込みデータを書き込む回路とが設けられる。

【００１９】

【作用】いずれかのドライブに障害が発生し、この障害が発生したドライブを正常なドライブに交換する際に、基板をマザーボードのコネクタから外しても、他の基板内のドライブのデータを用いて、取り外した基板のドライブのデータを回復できる。したがって、障害ドライブを含む基板をマザーボードのコネクタから外し、障害ドライブを正常ドライブに交換する作業を行っている間であっても、ＣＰＵからの読み出しや書き込み処理は中断せずにすむ。

【００２０】

【実施例】以下、図面を用いて本発明の実施例を説明する。

【００２１】（実施例１）

（１）装置の概要

図１は、本発明に係るディスクアレイシステムの第１の実施例の全体構成を示す。この図において、ディスクアレイシステムは、複数のドライブ５とそれらの動作を制御するアレイコントローラ２とからなる。

【００２２】１は外部からデータの書き込みおよび読み出しを指令するＣＰＵである。ＣＰＵ１とアレイコントローラ２とは、複数本のチャネルパス７により接続される。アレイコントローラ２には、チャネルパス７を選択するチャネルパスディレクタ８を分散して搭載した、複数、例えば４つのチャネルパスディレクタボード８Ａと、ＣＰＵ１とのデータ転送制御を行うためのチャネルインターフェース回路３をそれぞれ搭載した複数、例えば４つのチャネルパスインターフェースボード３Ａと、バッテリバックアップ等により不揮発化された半導体メモリであるキャッシュメモリ４を分散して保持する複数、例えば４つのキャッシュメモリボード４Ａと、これらのキャッシュメモリボード４Ａとチャネルインターフェースボード３Ａとを接続するバス３８とが設けられている。このバス３８に接続して、それぞれ複数のドライブ５とドライブインターフェース回路２０をそれぞれ保持する複数、例えば、４つのドライブボード５Ａが設けられている。

【００２３】図２（ａ）は、図１の装置に使用される複数のボードの配置を示す図である。各種のボードは、図２（ａ）に示すように、共通のマザーボード３７に搭載されている。図２（ａ）において、各ボード上に図示された四辺形３７０は、各ボードに搭載されたＬＳＩチップを示す。ただし、ドライブ用のボード５Ａの上に図示された四辺形の一部３７０はＬＳＩチップであるが残りの大部分５は、そのボードに搭載されたドライブを示す。このマザーボード３７上には、図２（ｂ）に示す、各ボード用のコネクタ群３０Ａ，４０Ａ，５０Ａ，８０Ａが設けられ、各ボードはこのようなコネクタによりマザーボードに対して着脱自在に保持されている。さらに、各ボードは互いにほぼ平行に配置されている。

【００２４】さらにマザーボード表面には、チャネルインターフェース用コネクタ群３０Ａ、キャッシュメモリ用コネクタ群４０Ａ、ドライブボード用のコネクタ群５０Ａを接続するように、前述のバス３８が形成されている。さらに、チャネルディレクタ用のコネクタ群８０Ａとチャネルインターフェース用のコネクタ群３０Ａとを接続する信号配線１２（図３）も設けられている。

【００２５】本実施例の特徴は、何れかのドライブに障害が発生したとき、そのドライブを搭載したドライブボードをマザーボード３７に接続したまま障害ドライブをそのドライブボードから分離するのではなく、そのドライブボードを、正常なドライブを搭載したままマザーボード３７から分離し、その障害ドライブを正常な交替ドライブにより置換した後、そのドライブボードをマザーボード３７に再度接続することを可能にした点にある。

【００２６】したがって、何れかのドライブに障害が発生したときに、そのドライブを搭載したドライブボードを抜き差しすればよいので、各ドライブボード間の間隔を小さくすることができる。しかも、複数のドライブボードに分散して搭載された複数のドライブでもって一つの論理グループが構成されている。これにより、何れかのドライブボードがマザーボードから分離された状態でも、その分離されたドライブボード上に搭載されたドライブに保持されたデータを、他のドライブボードに保持された複数のドライブに保持されたデータおよびパリティデータから回復可能になっている。

【００２７】さて、論理グループ９は、複数（本実施例では４枚）のボード間に渡るｍ台（本実施例では４台）のドライブ５により構成される。論理グループ９は、障害回復単位であり、この論理グループ９内のドライブ５は、ｍ－１個のデータとそれらから生成したパリティデータとからなる誤り訂正データグループ（すなわち、パ

リティグループ）を保持する。

【００２８】なお、本実施例では、チャネルインターフェースボード３Ａ、ドライブボード５Ａの数を各々４枚としたが、本発明はこれに限定されるものではない。また、ボードに装着されるドライブ５の数、および論理グループ９を構成するドライブ５の数も、本実施例に限定されず、任意の数としてよい。

【００２９】図３は、図１の各ボード内の回路の概略構成を示す。

【００３０】チャネルパスディレクタ８は、各チャネルパス７に接続されたインターフェースアダプタ（ＩＦＡｄｐ）１０、およびチャネルパススイッチ１１を備えている。このチャネルパスディレクタ８は、実際は４つのボードに分散して配置されるが、図では簡単化のために、この回路の全体を示してある。

【００３１】チャネルインタフェース回路３は、ＣＰＵ１から転送されてきたデータに対しプロトコル変換と転送速度を調整するチャネルインターフェース（ＣＨ　ＩＦ）１３、データの転送制御を行なうデータ制御回路（ＤＣＣ）１４、ＣＰＵ１から送られたデータをキャッシュメモリ４へ書き込むのを制御するチャネル側キャッシュアダプタ（Ｃ　Ａｄｐ）１５、このデータを分割して得られる複数のサブデータからパリティを生成するパリティ生成回路１８、およびこれらの回路を制御するマイクロプロセッサ（ＭＰ）１７を備えている。チャネルインタフェース１３は、データ線１２によりチャネルパススイッチ１１と接続されている。キャッシュアダプタ１５は、バス３８の制御と、ＣＰＵ１、キャッシュ４間のデータ転送制御を行なう回路である。

【００３２】ドライブボード５Ａには、ドライブ側キャッシュアダプタ（Ｃ　Ａｄｐ）１９、およびドライブインターフェース回路（Ｄｒｉｖｅ　ＩＦ）２０が設けられている。１６は、チャネルパススイッチ１１、チャネルインタフェース１３、データ制御回路１４、チャネル側キャッシュアダプタ１５、マイクロプロセッサ１７、キャッシュメモリ４、ドライブ側キャッシュアダプタ１９、およびドライブインターフェース回路２０を接続する制御信号線である。バス３８は、制御情報を転送する制御バスと、データを転送するデータバスからなる。

【００３３】なお、図１～図３においてボード５０Ａは後の実施例２で使用するもので、予備のドライブ５を搭載したボードである。

【００３４】（２）障害ドライブがないときの装置動作の概要

次に、図３を参照して、ディスクアレイシステムの内部動作を説明する。このときの動作は、基本的には、ＲＡＩＤ３の原理にしたがった動作である。

【００３５】ＣＰＵ１より発行された読み出しまたは書き込みコマンドは、チャネルパス７を通って、アレイコントローラ２のチャネルパスディレクタ８に入力する。

【００３６】ＣＰＵ１からコマンドが発行されると、アレイコントローラ２内のチャネルパスディレクタ８によりコマンドの受付が可能かどうか判断する。ＣＰＵ１からアレイコントローラ２に送られてきたコマンドはインターフェースアダプタ（ＩＦＡｄｐ）１０により取り込まれ、マイクロプロセッサ（ＭＰ）１７はコマンドの受付け処理を行なう。本実施例では、基本的に、従来のレベル３のＲＡＩＤによる方法にしたがってデータの書き込みおよび読み出しを行う。

【００３７】図６は、ＲＡＩＤレベル３によるデータの書き込みの様子を示す説明図である。図６において、ＢＡＤＲｉ（ｉ＝１，２，３または４）は、それぞれ、４つのドライブボード５Ａのボードアドレスを示す。ＳＤ＃ｉ（ｉ＝１，２，３または４）は、各ボード内の１つのドライブを特定するドライブアドレスを示す。

【００３８】以下、説明の便宜のため、例えばボードアドレスがＢＡＤＲ１であるボードは、ボードＢＡＤＲ１と呼ぶ。また、ボード内のドライブアドレスがＳＤ＃１のドライブは、ドライブＳＤ＃１と呼ぶ。

【００３９】ボードＢＡＤＲ１のドライブＳＤ＃１、ボードＢＡＤＲ２のドライブＳＤ＃１、ボードＢＡＤＲ３のドライブＳＤ＃１、およびボードＢＡＤＲ４のドライブＳＤ＃１により、論理グループが構成されているものとする。また、ボードＢＡＤＲ４のドライブは、パリティデータの格納に用いるものとし、他のボードＢＡＤＲ１からＢＡＤＲ３内のドライブＳＤ＃１は、データの格納に用いるものとする。

【００４０】ＣＰＵ１からのコマンドがデータの書き込みコマンドであり、ＣＰＵ１から転送された書き込みデータが、図６のＤ＃１であるとする。Ｄ＃１は、その書き込みデータを特定する論理アドレスである。以下、論理アドレスＤ＃１で特定されるデータをデータＤ＃１と呼ぶ。

【００４１】本実施例では、ＣＰＵ１から書き込みまたは読み出すデータのデータ長は常に４ＫＢの大きさの固定長データであるとする。したがって、ＣＰＵ１からの書き込みデータＤ＃１も４ＫＢである。なお、このデータ長により本発明が限定されることは無い。

【００４２】ＣＰＵ１からの書き込みデータＤ＃１は、図１のキャッシュメモリ４に保持される。図６に示すように、ＲＡＩＤ３ではこの保持された書き込みデータＤ＃１を、データの先頭から１バイトずつ、論理グループを構成する４台のドライブのうちデータ格納用の３台のドライブ（ＢＡＤＲ１，ＢＡＤＲ２，ＢＡＤＲ３のそれぞれのＳＤ＃１）に振り分けて分割する。分割したデータをサブデータと呼び、それぞれＤ＃１－１，Ｄ＃１－２，Ｄ＃１－３というサブデータ名を付ける。なお、サブデータ名がＤ＃１－１であるサブデータを、サブデータＤ＃１－１と呼ぶ。

【００４３】図６から分かるように、サブデータＤ＃１

－１はデータＤ＃１の第１バイト目、第４バイト目、第７バイト目、…からなり、サブデータＤ＃１－２はデータＤ＃１の第２バイト目、第５バイト目、第８バイト目、…からなり、サブデータＤ＃１－３はデータＤ＃１の第３バイト目、第６バイト目、第９バイト目、…からなる。

【００４４】サブデータに分割した後、これらのサブデータからパリティを作成する。このパリティは、サブデータにおいて各々対応するバイトに対して作成する。すなわち、サブデータＤ＃１－１の第１番目のバイトとＤ＃１－２の第１番目のバイトとＤ＃１－３の第１番目のバイトとからパリティを求め、サブデータＤ＃１－１の第２番目のバイトとＤ＃１－２の第２番目のバイトとＤ＃１－３の第２番目のバイトとからパリティを求め、…というようにパリティを生成する。生成したパリティを順に並べたパリティデータをＰ＃１と呼ぶ。

【００４５】これらのサブデータＤ＃１－１，Ｄ＃１－２，Ｄ＃１－３およびパリティデータＰ＃１を、論理グループを構成する４台のドライブに、分散して書き込む。ここでは、ボードＢＡＤＲ１のドライブＳＤ＃１にサブデータＤ＃１－１を書き込み、ボードＢＡＤＲ２のドライブＳＤ＃１にサブデータＤ＃１－２を書き込み、ボードＢＡＤＲ３のドライブＳＤ＃１にサブデータＤ＃１－３を書き込み、ボードＢＡＤＲ４のドライブＳＤ＃１にパリティデータＰ＃１を書き込むこととなる。

【００４６】以上のようにして、ＣＰＵ１からのデータの書き込みが行われる。

【００４７】ＣＰＵ１からデータの読み出しが要求されたときには、その要求されたデータを保持する論理グループから、そのデータを構成するサブデータを読み出し、それらのサブデータを結合して、キャッシュメモリ４を介してＣＰＵ１に送る。

【００４８】（３）アドレス変換用テーブル

以上のデータの書き込みおよび読み出しに必要なアドレス変換テーブルの詳細を以下に説明する。

【００４９】本実施例では、ディスクアレイシステムを構成するドライブ５としてＳＣＳＩインターフェースのドライブを使用する。また、本実施例では、ＣＰＵ１はディスクアレイを意識せずに書き込みコマンドや読み出しコマンドを発行するものとする。すなわち、ＣＰＵ１は、書き込みコマンドや読み出しコマンドを発行する相手であるディスク装置が従来方式のものかディスクアレイ装置であるかは意識していない。これは、現在主流のディスク装置が従来方式のものであり、ＯＳ（オペレーティングシステム）はディスクアレイを意識したものになっていないからである。

【００５０】このため、本実施例では、ＣＰＵ１では従来方式のドライブに対しデータを読み出しまたは書き込みを行っているものとして処理を行い、ディスクアレイシステムにおいて独自にディスクアレイで処理を行っている。したがって、ＣＰＵ１は従来のインターフェースで書き込みや読み出しのコマンドを発行する。

【００５１】具体的には、ＣＰＵ１からアレイコントローラ２にデータの書き込みを指示する場合、ＣＰＵ１は、書き込みデータに論理アドレス（データ名またはデータ番号）を付して転送する。アレイコントローラ２では、図６で説明したように、この書き込みデータをサブデータに分割し、パリティデータを作成し、それらのサブデータおよびパリティデータを論理グループ９の各ドライブに格納する。この際、それらのサブデータおよびパリティデータには、アレイコントローラ２において処理するアドレス（図６で説明したＤ＃１－１～Ｄ＃１－３やＰ＃１など）がつけられる。ＣＰＵ１から与えられる論理アドレスとこのアレイコントローラ２におけるアドレスとの対応をとるテーブルが、以下で説明するアドレス変換用テーブルである。

【００５２】ＣＰＵ１からアレイコントローラ２にデータの読み出しを指示する場合、ＣＰＵ１は、論理アドレスを与えてアレイコントローラ２に読み出しを指示する。アレイコントローラ２は、以下で説明するアドレス変換用テーブルを利用して論理アドレスをディスクアレイシステム内のアドレスに変換し、そのアドレスでサブデータを読み出して結合してＣＰＵ１に渡す。

【００５３】アドレス変換用テーブルは、論理グループテーブル２１とアドレステーブル３１の２個のテーブルにより構成される。

【００５４】図４は、論理グループテーブル２１の構造を示す。論理アドレス（論理Ａｄｄｒ）２２は、ＣＰＵ１から指定されるアドレスであるところの論理アドレス（データ名またはデータ番号）である。サブデータ名２４は、対応する論理アドレス２２のデータを分割して得たサブデータの名称である。パリティ名２５は、対応するサブデータから作成されたパリティデータの名称である。論理グループ番号（Ｎｏ．）２３は、これらのサブデータおよびパリティデータが実際に格納されているまたは格納する論理グループ９の番号である。

【００５５】データを全く書き込んでいない初期状態において、論理グループ番号（Ｎｏ．）２３の欄にはあらかじめ設定されている論理グループの番号が初期値として設定されている。初期状態では、論理アドレス２２、サブデータ名２４、およびパリティ名２５は、すべて空欄で空きを示すことになる。

【００５６】図５は、アドレステーブル３１の構造を示す。アドレステーブル３１は、論理グループのどのドライブのどの位置にデータが実際に書き込まれているか（または書き込むか）、その詳細なアドレス情報を保持する。

【００５７】図５のアドレステーブル３１において、論理アドレス（論理Ａｄｄｒ）２２は、図４の論理グループテーブル２１の論理アドレス２２と同じであり、ＣＰ

Ｕ１から指定されるアドレスであるところの論理アドレス（データ名またはデータ番号）である。また、論理グループ番号（Ｎｏ．）２３も図４の論理グループテーブル２１の論理グループ番号２３と同じである。３０は、対応する論理アドレスのデータが実際に書き込まれている（または書き込む）ＳＣＳＩドライブのアドレス情報を示す。

【００５８】ＳＣＳＩドライブアドレス情報３０は、４列のアドレス情報からなる。各列は、ボードアドレス（Ａｄｄｒ）２７、ドライブ番号（Ｄｒｉｖｅ　Ｎｏ）２８、障害フラグ２９およびボード抜去フラグ１００からなる。この内、ボード抜去フラグ１００は、そのボードアドレスのボードがマザーボードから抜き去られているか否かを示す。３２は、実際にデータが各ドライブのどこに書き込まれているか（または書き込むか）を示すドライブ内アドレスである。論理アドレス２２に対応するボードアドレス２７、ドライブ番号２８、およびドライブ内アドレス３２により、その論理アドレス２２のデータ（サブデータとパリティデータ）が書き込まれている（または書き込む）位置が、どのボードのどのドライブのどの位置かを知ることができる。

【００５９】各ボードに対するボード抜去フラグ１００は、そのボードをマザーボード３７から引き抜いたときに１にセットされる。ユーザは何れかのボードを引き抜いたとき、ユーザがＣＰＵ１を介して、このボードを引き抜いたことを示すコマンドをアレイコントローラ２に発行するようになっている。ＭＰ１７はこのコマンドに応答して、そのコマンドが示すボードに対応するボード抜去フラグ１００を１にセットする。また、後に、この引き抜かれたボードが、マザーボードに再度挿入されたときには、同様にしてユーザがＣＰＵ１からのコマンドにより、このボードが搭載されたことをアレイコントローラ２に通知する。アレイコントローラ２は、このときに、後に述べるような回復処理をした後に、このボードに対するボード抜去フラグ１００をリセットする。

【００６０】障害フラグ２９は対応するボード内の対応するドライブに障害が発生したため、読み出しまたは書き込みができない場合オン（１）、正常の場合オフ（０）となり、ドライブの障害の有無を判定することが可能となる。

【００６１】ＳＣＳＩドライブアドレス情報３０の４つの列は、図４の論理グループテーブル２１のサブデータ名２４およびパリティ名２５の列に順に対応している。

【００６２】例えば、図４の論理グループテーブル２１によれば、論理アドレスＤ＃１に対応する論理グループ番号２３はＬＧ＃１、サブデータ名２４とパリティ名２５は順にＤ＃１－１、Ｄ＃１－２、Ｄ＃１－３、Ｐ＃１である。したがって、論理アドレスＤ＃１のデータは、３つのサブデータＤ＃１－１、Ｄ＃１－２、Ｄ＃１－３に分割されてパリティデータＰ＃１が作成され、それらが論理グループＬＧ＃１の各ドライブに格納されていることが分かる。

【００６３】さらに、図５のアドレステーブル３１により、各サブデータおよびパリティデータが実際に格納されている位置が分かる。具体的には、サブデータＤ＃１－１はボードアドレス２７がＢＡＤＲ１のボードのドライブ番号２８がＳＤ＃１のドライブ５に格納され、サブデータＤ＃１－２はボードアドレス２７がＢＡＤＲ２のボードのドライブ番号２８がＳＤ＃１のドライブ５に格納され、サブデータＤ＃１－３はボードアドレス２７がＢＡＤＲ３のボードのドライブ番号２８がＳＤ＃１のドライブ５に格納され、これらのサブデータにより作成されたパリティデータＰ＃１はボードアドレス２７がＢＡＤＲ４のボードのドライブ番号２８がＳＤ＃１のドライブ５に格納されていることが分かる。さらに、これらのサブデータおよびパリティは、各々のドライブ５内のドライブ内アドレス３２がＳＡＤＲ１の位置に格納されていることが分かる。

【００６４】データを全く書き込んでいない初期状態において、図５のアドレステーブル３１の論理グループ番号２３の欄には、あらかじめ設定されている論理グループの番号が初期値として設定されている。また、初期状態において、ボードアドレス（Ａｄｄｒ）２７、ドライブ番号（Ｄｒｉｖｅ　Ｎｏ）２８、およびドライブ内アドレス３２の欄もあらかじめ設定されている。初期状態では、論理アドレス２２がすべて空欄となっており、該当する位置が空き領域であることになる。

【００６５】上記のアドレス変換用テーブル（論理グループテーブル２１とアドレステーブル３１）は、動作時には、アレイコントローラ内のキャッシュメモリ４内の適当な領域に格納されている。これらのアドレス変換用テーブルは、システムの電源をオンしたときに、ＭＰ１７により論理グループ９内のある特定のドライブ５からキャッシュメモリ４に自動的に読み込まれる。一方、電源をオフするときは、ＭＰ１７によりキャッシュメモリ４内のアドレス変換用テーブルが、元のドライブ５内の所定の場所に自動的に格納される。このような電源オン／オフの際のアドレス変換用テーブルのロードとストアはＣＰＵ１の関与なく行なわれる。

【００６６】なお、本実施例では、同じドライブ内アドレス３２を有する複数の領域にデータを並列に格納するため、論理グループ９内の各ドライブ５の回転をすべて同期させる方が望ましい。

【００６７】本実施例でのデータの書き込みと読み出しは、このようなアドレステーブル３１、および論理グループテーブル２１を使用して行なわれる。

【００６８】（４）ドライブに障害が発生したときの装置動作

本実施例によれば、図６からも分かるように、複数のボード（４枚のボードＢＡＤＲ１，ＢＡＤＲ２，ＢＡＤＲ

３，ＢＡＤＲ４）にまたがる複数のドライブで論理グループ９を構成するようにしており、図４，５のアドレス変換用テーブルにおいては各ドライブのアドレスにボードアドレスを付加して、どのボードのどのドライブで論理グループを構成しているのかが明確になるようにしている。

【００６９】本実施例では、何れかのドライブ、例えばＢＡＤＲ１上のドライブに読み出しまたは書き込み処理を行なおうとしたが、処理が行なえず、何回かリトライを行なっても処理できない場合、ＭＰ１７は、ボードＢＡＤＲ１に障害が発生したと判断し、アドレステーブル３１のボードＢＡＤＲ１のドライブＳＤ＃１に対応する障害フラグ２９をオン（１）にする（図５）。このように、このドライブに障害が発生したとき、そのドライブを搭載するボードＢＡＤＲ１がマザーボード３７に搭載した状態でも、このボードをマザーボード３７から取り外して修理に供している状態でも、このディスクアレイシステムは引続き使用できるようになっている。

【００７０】すなわち、ＣＰＵ１からこの各ボードＢＡＤＲ１－４上の各ドライブＳＤ＃１に保持されたデータに対して読み出し要求が発行された場合、ＭＰ１７は、アドレステーブル３１を調べ、アドレス変換用テーブルに関連して先に述べたような手順で、ＣＰＵ１から指定された論理アドレスをディスクアレイシステム内のアドレスに変換する。このとき、その変換後のアドレスが割り当てられているドライブ、今の例ではボードＢＡＤＲ１上のドライブＳＤ＃１に対応する障害フラグ２９と、このドライブが搭載されているボードに対応するボード抜去フラグ１００を調べる。このドライブＳＤ＃１に対して障害フラグ２９がオン（１）で、このドライブが搭載されているボードＢＡＤＲ１のボード抜去フラグ１００がオフ（０）の場合は、ＭＰ１７は、このドライブに障害が発生しており、しかも、このドライブが搭載されているボードＢＡＤＲ１はマザーボード３７に接続されていることを認識する。

【００７１】このようにドライブの障害をＭＰ１７が認識すると、ＭＰ１７は、このドライブが属する論理グループのデータとパリティデータとから、ボードＢＡＤＲ１上のドライブＳＤ＃１のデータを回復する。すなわち、他のボードＢＡＤＲ２－４上の他の複数のドライブＳＤ＃１から複数のサブデータおよびパリティデータを読み出し、これらのサブデータおよびパリティデータから障害ドライブＳＤ＃１のサブデータを、ＭＰ１７はパリティ生成回路１８を使用して回復し、先に読み出した複数のサブデータと結合して、要求されたデータを生成し、その生成したデータをＣＰＵ１に送る。障害ドライブＳＤ＃１内のデータの回復をパリティ生成回路１８により行なわせることは、それ自体は公知である。

【００７２】このように、パリティ生成回路（ＰＧ）１８は、データをドライブへ書込む場合は、サブデータからパリティを生成する回路であるが、サブデータとパリティからサブデータを回復する場合はデータ回復回路として使用される。

【００７３】一方、ドライブＳＤ＃１の搭載されているボードＢＡＤＲ１のボード抜去フラグ１００がオン（１）の場合は、ＭＰ１７はボードＢＡＤＲ１がマザーボード３７から抜かれていると認識する。このようにＭＰ１７が認識すると、ＭＰ１７は、このボードＢＡＤＲ１に搭載されている全ドライブは、障害がなくても読み出し処理ができない。このため、先に述べた障害が発生したドライブＳＤ＃１での処理と同様にして、それらのドライブがアクセスされたときに、そのアクセスされたドライブ内のデータを、そのドライブと同じ論理グループに属する他のドライブのデータから回復する。

【００７４】また、ＣＰＵ１からドライブＳＤ＃１に保持されたデータを更新する書き込み要求がきた場合、ＭＰ１７は、読み出し要求での処理と同様に、アドレステーブルを調べ、アドレス変換用テーブルのところで述べたような手順で、ＣＰＵ１から指定された論理アドレスをディスクアレイシステム内のアドレスに変換する。このとき、該当するドライブに対応する障害フラグ２９と、この当該ドライブが搭載されているボードに対応するボード抜去フラグ１００を調べる。

【００７５】ドライブＳＤ＃１に対して障害フラグ２９がオン（１）で、このドライブＳＤ＃１が搭載されているボードＢＡＤＲ１のボード抜去フラグ１００がオフ（０）の場合は、ＭＰ１７は、ドライブＳＤ＃１に障害が発生しており、しかも、このドライブＳＤ＃１が搭載されているボードＢＡＤＲ１はマザーボード３７に接続されていると認識する。このように、ＭＰ１７が認識すると、ＭＰ１７は、この書き込みデータはキャッシュメモリ４内に保持しておく。

【００７６】後に、このボードＢＡＤＲ１の障害が発生したドライブＳＤ＃１の修理が完了した時点で、このボードＢＡＤＲ１がマザーボード３７に再度組み込まれたときに、この障害が発生したドライブに代わる正常なドライブ、今の例ではボードＢＡＤＲ１の新たなドライブＳＤ＃１と同じ論理グループに属する他のドライブのすべての誤り訂正データ群に属するサブデータおよびパリティデータを読み出し、それらのサブデータおよびパリティデータからその障害が発生したドライブ内の全データを回復し、この障害が発生したドライブを置換した正常な交替ドライブに格納する。

【００７７】その後、キャッシュメモリ４に保持されていた複数の書き込みデータに対して、書き込み動作を実行して、それぞれの書き込むデータを、上記ボードＢＡＤＲ１上の、上記交替ドライブあるいは他の障害が発生しなかった正常なドライブに書き込む。この書き込みを行なうときには、ＲＡＩＤ３による書き込みを行なうために、この保持された各書き込みデータを複数のサブデ

ータに分割し、それらからパリティデータを生成し、分割により得られた複数のサブデータの一つを上記ボード上の一つのドライブに書き込み、他の複数のサブデータと生成されたパリティデータを他のボード上の複数のドライブに書き込む。

【００７８】一方、ドライブＳＤ＃１が搭載されているボードＢＡＤＲ１のボード抜去フラグ１００がオン（１）の場合は、ＭＰ１７はボードＢＡＤＲ１がマザーボード３７からぬかれていると認識する。このようにＭＰ１７が認識すると、ＭＰ１７は、このボードＢＡＤＲ１に搭載されている全ドライブは、障害の有無にかかわらず書き込み処理ができないため、先に述べた障害が発生したドライブＳＤ＃１での処理と同様にドライブＳＤ＃１に書き込むデータをキャッシュメモリ４内に保持しておく。

【００７９】後に、この障害が発生したドライブを正常な交替ドライブで置換した後のドライブボードＢＡＤＲ１がマザーボード３７に再度組み込まれたときに、この障害が発生したドライブを置換した交替ドライブと同じ論理グループに属する他のドライブのすべての誤り訂正データ群に属するサブデータおよびパリティデータを読み出し、それらのサブデータおよびパリティデータからその障害が発生したドライブ内の全データを回復して、その交替ドライブに書き込む。その後、キャッシュメモリ４に保持されていた書き込みデータに対して、書き込み動作を実行する。

【００８０】なお、本実施例では、分離されたボードＢＡＤＲ１内の障害が発生していないドライブも、このボードが分離されている間は、アクセスできなくなる。これらのドライブに対するデータの書き込みあるいは読み出しも、その障害ドライブを搭載したボードＢＡＤＲ１がマザーボード３７から分離されている場合に、その障害があるドライブに対して行なうのと同じように行なう必要がある。ただし、この分離されたボード上の正常なドライブに対して書き込み要求が発生した場合には、そのボードが後に再度マザーボードに接続された時点では、上記障害が発生したドライブを置換したドライブに対して行なうと説明した、そのドライブ内の全データの回復を、その正常なドライブに関して行なう必要はない。そのような正常なドライブは、そのボードを分離するまでのデータを正常に保持しているからである。

【００８１】このように新たに発生した書き込みデータをキャッシュメモリ４に保持するので、マザーボード３７から分離されていた間に発生した書き込みデータを、分離されたボード内の正常なドライブに書き込むことが容易に行ない得る。

【００８２】また、キャッシュメモリ４に書き込みデータを保持しているので、後にその保持されたデータを読み出す要求がＣＰＵ１から供給された場合、キャッシュメモリ４からその保持されたデータをＣＰＵ１に読み出しデータとして供給すればよいので、読み出しが実質的に高速になるという利点も有する。

【００８３】以上のごとくにして、本実施例では、障害発生により何れかのドライブボードが取り去られた後でも、そのドライブに関連するデータの読み出しおよび書き込み要求を引き続き処理することができる。

【００８４】（実施例１の変形例１）障害があるドライブを保持するボードを分離した状態で、ＣＰＵ１から供給された、そのボード上のドライブ障害があるドライブＳＤ＃１に対する書き込み要求（あるいはそのボード上の他の正常なドライブに対する書き込み要求）は、以下のように処理することも可能である。

【００８５】すなわち、この書き込みデータをサブデータに分割し、それからパリティを生成するところまでは、ボードＢＡＤＲ１が分離されていない場合と同様に行なう。さらに、ボードＢＡＤＲ１が分離されている場合でも、このボード上のドライブＳＤ＃１（あるいは他の正常なドライブ）が属する論理グループの内、この障害が発生したドライブ（あるいは正常なドライブ）を有するボードＢＡＤＲ１以外のボードＢＡＤＲ２－４内のドライブＳＤ＃１あるいは他のドライブに、それぞれのサブデータあるいは新たに生成されたパリティを書き込む。分離されたボードＢＡＤＲ１内のドライブＳＤ＃１あるいは他の正常なドライブに書き込むべきサブデータは捨てる。

【００８６】後に、このボードＢＡＤＲ１が、マザーボード３７に再度接続された時点で、実施例１で述べたように、この分離されたボードに保持されたすべてのドライブに対して、それぞれのドライブのデータの回復を行なう。

【００８７】この回復により、障害が発生したドライブ内の元のデータおよびそのボードを分離中に発生した書き込み要求に付随する書き込みデータを、このボードの、この障害ドライブを交替する新たなドライブに回復するとともに、このボードの他の正常なドライブにも、障害発生時点の前からそれらの正常なドライブが保持していたデータおよびそのボードの分離中に発生した新たな書き込み要求に付随する書き込みデータを回復できる。

【００８８】上記方法では分離されていたボードＢＡＤＲ１がマザーボード３７に再接続された際、ボードＢＡＤＲ１上の正常なドライブに対し全データの回復を行なう。この方法では、ボードＢＡＤＲ１が分離中に書き込み要求が発生していないデータも回復する。ボードＢＡＤＲ１が分離中に非常に多くの書き込み要求が発生し、正常なドライブ内の大部分のデータが更新されている場合は有効である。しかし、分離中に書き込み要求があまり発生しなかった正常なドライブに対しては、すでに保持されている書き込みデータばかりを回復することになり、回復処理時間が無駄である。

【００８９】そこで、分離中に書き込み要求があまり発生しなかった正常なドライブに対しては、すでに保持されていた書き込みデータは回復せず、分離中に上位装置から発行された書き込みデータのみを回復する。

【００９０】この変形例では、分離されたボード上のドライブの数が多いときには、それらのドライブのデータの全体を回復するのに時間がかかるが、実施例１のように、新たな書き込みデータをキャッシュメモリ４に保持しなくてもよく、それだけキャッシュメモリ４のサイズが小さくてもよいという利点がある。

【００９１】（実施例１の変形例２）変形例１と同様に、障害が発生したドライブを有するボードＢＡＤＲ１をマザーボードから分離した状態において、ＣＰＵ１から書き込み要求が供給されたときに、そのデータを複数のサブデータに分割し、これらの複数のサブデータからパリティデータを生成する。これらのサブデータの内、上記ボードＢＡＤＲ１内のドライブに書き込むべきサブデータ以外と、上記生成されたパリティデータを、上記ボードＢＡＤＲ１以外の、互いに異なる複数のボード上の複数のドライブに書き込む。実施例１の変形例１と異なり、これらのサブデータの一つが上記ボードＢＡＤＲ１上の、障害が発生していない正常なドライブに書き込むべきサブデータであるときには、そのサブデータをキャッシュメモリ４に保持する。後にボードＢＡＤＲ１が再度マザーボードに接続された時点で、実施例１の変形例１と同様に、ボードＢＡＤＲ１上の、障害ドライブを置換した交替ドライブが保持すべき全データを他のドライブのデータを利用して回復する。一方、ボードＢＡＤＲ１に保持された正常ドライブには、実施例１と同様に、キャッシュメモリ４に保持されたサブデータを書き込む。

【００９２】この方法によれば、キャッシュメモリに保持されるサブデータの量が実施例１より少なくなり、かつ、ボードＢＡＤＲ１がマザーボードに再度接続された時点で、ボードＢＡＤＲ１上の正常なドライブに対して、実施例１の変形例１で行なったデータの回復は行なう必要がない。

【００９３】（実施例１の変形例３）本実施例１では、バス３８を複数のドライブボードで共通で使用される高速なバスとしたが、各ドライブボードとアレイコントローラ２とを１対１で結ぶ、そのドライブボード専用のバスとしてもよい。このように専用のバスにすると、複数のドライブ５を同時に並列に動作させることが可能となる。

【００９４】（実施例２）本実施例では、図１あるいは図３に点線で示した予備ボード５０Ａがさらに付加されている点で実施例１と異なる。図２（ｂ）では、５００Ａはこの予備ボード用のコネクタを例示する。予備ボード５０Ａの構造は他のドライブボードと同じであり、そのボード上に他のボード上にあるドライブと同数のドライブが搭載され、予備ドライブとして使用される。本実施例では、各予備ドライブは、他の複数のボード上に搭載されたドライブにより形成されている複数の論理グループの一つに対応し、その対応する論理グループ内の何れかのドライブに障害が発生したとき、その障害が発生したドライブの代わりに使用される。以下では、この予備ボード以外のドライブボードの番号を、実施例１と同様に、ＢＡＤＲｉ（ｉ＝１，２，３または４）で表し、この予備ボードのボード番号をＢＡＤＲ５とする。

【００９５】この予備ドライブの存在に伴い、本実施例では論理グループテーブル２１として図７に示すものが、図３に示すものに代えて使用される。本実施例では、予備ドライブに格納するデータ名の欄２６が設けられている。アドレステーブル３１には、図５に代えて図８に示すものが使用される。このアドレステーブル３１は、各論理グループを構成するドライブの総数は、一つの予備ドライブを含む、総計５つのドライブからなる点で図５と異なる。

【００９６】本実施例における、ドライブに障害が発生したときの動作は以下の通りとなる。以下では、ボードＢＡＤＲ１上のドライブＳＤ＃１に障害が発生したときの装置動作を説明する。

【００９７】（Ａ）障害ドライブに関連する装置動作

（Ａ１）障害ドライブを有するボードがマザーボードに接続されている場合

（Ａ１ａ）障害ドライブに保持されたデータの読み出し障害ドライブを有するボードＢＡＤＲ１がマザーボード３７に接続された状態で、ＣＰＵ１からこの障害ドライブに保持されたサブデータを一部に含むデータに対して読み出し要求が発生されたとき、実施例１の場合と同様に、この障害ドライブと同じ論理グループ内に属する他のドライブのサブデータおよびパリティから障害ドライブ内のサブデータを回復する。今の例ではボードＢＡＤＲ２内のドライブＳＤ＃１、ボードＢＡＤＲ３内のドライブＳＤ＃１から二つのサブデータおよびボードＢＡＤＲ４内のドライブＳＤ＃１内のパリティデータとから、ボードＢＡＤＲ１の障害ドライブＳＤ＃１に保持されたサブデータを回復する。その回復されたサブデータを、他の読み出された二つのサブデータと結合して、ＣＰＵ１に送る。

【００９８】本実施例では、この後の動作が実施例１と異なる。すなわち、この回復されたサブデータを、この論理グループに対して設けられた予備ドライブ、今の例ではボードＢＡＤＲ５のドライブＳＤ＃１に書き込む。その後、このサブデータに関しては、この予備のドライブを障害の発生したドライブに代えて使用する。すなわち、この後、上の読み出し要求で指定されたデータに対して、ＣＰＵ１から再度読み出し要求が発行されたときには、あるいはこのデータに対して書き換え要求が出されたとき、この予備のドライブを含む４つの正常なドラ

イブに対して、読み出し要求あるいは書き込み要求を実行する。

【００９９】以上の動作のために、予備ボードＢＡＤＲ５内の予備ドライブＳＤ＃１内のアドレスＳＡＤＲ１の位置を、障害ドライブの代わりに使用するように、アドレステーブル３１Ａを書き換えるのは言うまでもない。

【０１００】以上の動作から分かるように、この予備ドライブを使用することにより、一度アクセスされたデータをその後再度アクセスするときには、実施例１で必要であったサブデータの回復が不要になる。さらにそのデータを書き換えるときには、実施例１のごとく、書き込みデータをキャッシュメモリ４に保持しておく必要がなく、その書き込み要求を実行できる。

【０１０１】（Ａ１ｂ）障害ドライブに保持されたデータに対する書き込み動作

障害ドライブを有するボードがマザーボード３７に搭載された状態で、この障害ドライブが保持するサブデータを含むデータに対して，ＣＰＵ１からデータ書き込み要求が発行された場合、実施例１と異なり、この書き込み要求を即実行する。

【０１０２】すなわち、通常の書き込み動作と同様に、この書き込みデータを分割して３つのサブデータを得、それらのサブデータからパリティデータを生成する。

【０１０３】通常の書き込み動作と異なるのは、これらのサブデータの内、障害が発生した、ボードＢＡＤＲ１内のドライブＳＤ＃１に書き込むべきサブデータは、予備ボードＢＡＤＲ５内のドライブＳＤ＃１に書き込むことである。

【０１０４】もちろん、この動作のために、アドレステーブル３１Ａを予備ドライブを使用するように書き換える。

【０１０５】（Ａ２）障害ドライブを有するボードがマザーボードから分離接続されている状態での装置動作

この場合の障害ドライブに保持されたデータの読み出しおよびデータ書き込み動作は、障害ドライブがマザーボードに接続されている場合の、データ読み出し動作（Ａ１ａ）およびデータ書き込み動作（Ａ１ｂ）と同じである。

【０１０６】（Ａ３）障害ドライブを有するボードをマザーボードに復元した後の装置動作

障害ドライブが正常なドライブにより置換された後、障害ドライブを有するボードＢＡＤＲ１がマザーボード３７に再度搭載された後では、実施例１と同様に、元の障害ドライブが保持していたデータを回復し、障害ドライブを置換したドライブに書き込む。この回復に当たっては、予備ドライブ内に保持されているデータを使用する必要はない。なお、この方法によれば、障害が発生した後に、新たにＣＰＵ１から書き込まれたデータも回復される。

【０１０７】したがって、この障害ドライブの代わりに使用した、障害ドライブを有するボードＢＡＤＲ５上の予備ドライブは、この後は、同じ論理グループに属する他のドライブに障害が発生したときにその新たな障害ドライブの代わりに使用されることになる。

【０１０８】（Ｂ）障害ドライブを有するボード上の正常ドライブに関連する装置動作

（Ｂ１）障害ドライブを有するボードがマザーボードに接続されている場合

実施例１と同様に、障害ドライブを有するボード上の正常ドライブはそのまま引き続き使用される。

【０１０９】（Ｂ２）障害ドライブを有するボードがマザーボードから分離されている場合

この場合には、障害ドライブを有するボード上の正常なドライブはアクセスできない。したがって、前述の障害ドライブの動作（Ａ１）で説明された動作を行なう。

【０１１０】（Ｂ３）障害ドライブを有するボードをマザーボードに復元した後の装置動作

この場合、分離されていた障害ドライブを有するボード上の正常なドライブが保持していた全データの内、障害ドライブを有するボードがマザーボードから分離した後に、この正常ドライブに書き込み要求が発生している場合、その正常なドライブの元のデータは使用できない。この書き込み要求に伴って生成された、この正常なドライブに書き込むべきデータは、前述した通り、この正常なドライブの予備のドライブに書き込まれている。したがって、障害ドライブを有するボードがマザーボード３７に再度接続された後は、この正常なドライブに対する予備ドライブ内に保持されている有効なデータを、障害ドライブを有する、再度接続されたボード内の正常なドライブに移動すればよい。こうして、分離された障害ドライブを有するボード上の正常なドライブに対するデータの回復は完了する。

【０１１１】なお、障害ドライブを有するボードが分離されている間に、この正常なドライブに保持されているデータに対して読み出し要求がＣＰＵ１より発行された場合、すでに述べたごとく、この正常なドライブに保持されているデータが回復され、このドライブに対する予備ドライブに格納されている。したがって、上記データの移動により、このような読み出し要求に伴い回復されたデータも、予備ドライブから正常なドライブに移動されることになる。しかし、このように読み出し要求によって回復されたデータは、もともと正常なドライブに保持されているデータと同じである。したがって、上記移動のときに、このようなデータを移動しても、移動完了後の、正常なドライブのデータは正常な値を有する。

【０１１２】このように、本実施例では、実施例１と異なり、予備のドライブが設けられ、何れかのドライブボードがマザーボード３７から分離されている間、ＣＰＵ１から供給された読み出し要求に応答して回復されたデータが、この予備のドライブに保持される。したがっ

て、そのようなデータがその後ＣＰＵ１から要求されたときに、すぐに利用可能になるという利点を有する。もちろん、キャッシュメモリ４に書き込みデータを保持する実施例１と異なり、本実施例は、書き込みデータを予備のドライブに保持するので、より多くの書き込みデータを保持できるという利点がある。

【０１１３】（実施例２の変形例）実施例２に対して、障害ドライブを有するボードをマザーボードに復元した後のデータの回復に関連して、次のようないろいろの変形が可能である。

【０１１４】（１）障害ドライブのデータの回復に関連する変形

（１ａ）上述の実施例２では、障害ドライブを置換したドライブが保持すべき全データを、それが属する論理グループ内の他のドライブ内の複数のサブデータとパリティデータとを使用して回復し、その置換したドライブに格納した。

【０１１５】障害ドライブの予備のドライブには、障害ドライブが保持すべきデータの内の一部のデータがすでに保持されている。したがって、実施例２における、障害ドライブを置換したドライブへのデータの回復に当たり、この予備のドライブに保持されているデータ以外のデータは、それらを回復して、その置換したドライブに格納するが、予備のドライブにすでに保持されているデータは、その予備のドライブからその置換したドライブに移動すれば、実施例２より高速にデータの回復を完了できる。この方法は、予備のドライブに格納されているデータ量が多いときに有効である。

【０１１６】（１ｂ）実施例２では、障害ドライブを有するボードが復帰したとき、もとの障害ドライブを置換したドライブに、もとの障害ドライブが保持すべきであったデータを回復し、その障害ドライブに対して使用した予備ドライブは、新たに予備ドライブとして解放した。しかし、この方法に代えて、障害ドライブが保持すべきデータを回復して、予備のドライブに書き込み、この予備のドライブを引き続き正常なドライブとして使用することもできる。この際、もとの障害ドライブが保持すべき全データの内、この予備ドライブにまだ保持されていないデータを選択的に回復して、この予備のドライブに書き込むことが回復時間を短縮する上で望ましい。

【０１１７】この方法は、もとの障害ドライブが保持すべきデータの内、かなり大きな割合のデータがすでに予備ドライブに保持されている場合に有効である。

【０１１８】（１ｃ）実施例２では、上位装置から障害ドライブ内のデータに対する読み出し要求が供給されたときに、その供給を契機として、そのデータを回復し、さらに、この回復後のデータを予備ドライブに格納した。しかし、この予備ドライブへの格納を省略できる。すなわち、予備ドライブには、障害があるドライブを搭載したボードがマザーボードから分離されている間に上位装置から供給された書き込みデータだけを保持する。もちろん、このボードがマザーボードに再度接続された後は、障害ドライブのデータを回復して予備ドライブに格納し、その予備ドライブをその障害ドライブの代わりに使用してもよい。

【０１１９】（１ｄ）何れかのドライブに障害が発生したときに、直ちに上位装置からディスクアレイへのアクセスを禁止し、その障害ドライブの全データを回復し、予備ドライブに格納し、その後は、この予備ドライブを障害が発生したドライブの代わりに使用してもよい。

【０１２０】（２）障害ドライブを有するボード上の正常ドライブに関連する変形例

（２ａ）実施例２では、障害ドライブを有するボードをマザーボードに復元した後、障害ドライブを有するボード上の正常なドライブに対する予備のドライブにあるデータをすべて、その正常なドライブに移動した。しかし、この予備のドライブにあるデータには、障害ドライブを有するボードが分離されている間にＣＰＵ１から発生された、その正常なドライブのデータに対する読み出し要求を契機として回復されたデータも含まれている。このようなデータはすでにその正常なドライブに保持されている、その障害ドライブを有するボードを分離される前のデータと同じ値を有する。したがって、実施例２における、予備ドライブから、正常ドライブへのデータの移動のときに、このような読み出しを契機として回復されたデータは移動しないようにすることが、移動を速く完了する上で望ましい。

【０１２１】このためには、予備のドライブにデータが書き込まれたとき、そのデータが書き込み要求を契機として書き込まれたデータか読み出し要求を契機として書き込まれたデータか否かを示す情報をアドレステーブルに格納し、その情報にしたがって、そのデータを予備ドライブから復帰されたボード上の正常ドライブに移動するか否かを制御すればよい。この方法は、予備ドライブに保持されたデータの内、ＣＰＵ１からの書き込み要求を契機として書き込まれたデータが、ＣＰＵからの読み出し要求を契機として書き込まれたデータより多いときに有効である。

【０１２２】（２ｂ）実施例２では、障害ドライブを有するボードが復帰したとき、このボード上の正常なドライブに、そのドライブの予備のドライブに書き込まれたデータを移動した。しかし、この方法に代えて、この正常な復帰したドライブが保持すべきデータを回復して、この予備のドライブに書き込み、この予備のドライブを引き続き正常なドライブとして使用することもできる。この際、この正常なドライブが保持すべき全データの内、この予備ドライブにまだ保持されていないデータを選択的に回復して、この予備のドライブに書き込むことが有効である。

【０１２３】この方法は、正常なドライブに保持すべき

データの内のかなりの割合のデータがすでに予備ドライブに保持されている場合に有効である。

【０１２４】（実施例３）実施例１、２はＲＡＩＤ３の制御を適用したディスクアレイシステムであるが、本実施例３は本願発明を実施例１の装置の制御にＲＡＩＤ５を採用した実施例である。したがって、本実施例は、図１に示すものと同じ装置構成により実現されるが、ＭＰ１７による制御が実施例１と異なる。したがって、本実施例でも図１などの回路を使用して本実施例を説明する。説明は実施例１と異なる点を主にする。図９は、この実施例で使用するアドレステーブルの一例を示す。

【０１２５】（ドライブ障害がないときの装置動作）このときの装置の動作は、基本的には、ＲＡＩＤ５の動作として周知の動作と同じであるが、ＲＡＩＤ５の動作の理解のための簡単にこのときの動作を説明する。

【０１２６】ＲＡＩＤ３では、一つの書き込み要求に付随した書き込みデータが複数のサブデータに分割され、それらのサブデータからパリティデータを生成する。生成された複数のサブデータとパリティデータを互いに異なるドライブに書き込む。ＲＡＩＤ５ではＲＡＩＤ３とは異なり、一つの書き込み要求に付随するデータは分割されない。複数の書き込み要求に付随する複数のデータに対してパリティデータが定義される。これらの複数のデータとそのパリティデータが一つのパリティグループを形成する。これらのデータとパリティデータとは互いに異なるドライブに書き込まれる。しかし、同一のパリティグループに属する複数のデータが揃ってからそれらのデータとパリティデータが複数のドライブに書き込まれるのではない。各書き込み要求がＣＰＵから供給されるごとに、そのデータが属するパリティグループに対してすでに生成されたパリティデータ（旧パリティデータ）とその書き込みデータとから新パリティデータを生成する。その書き込みデータと生成されたパリティデータたが二つのドライブに書き込まれる。読み出し要求がＣＰＵから供給されたときには、その読み出し要求で指定されるデータを何れか一つのドライブから読み出し、ＣＰＵ１に供給する。

【０１２７】すなわち、互いに異なる複数のドライブに属する複数の記憶位置を何れかのパリティグループ用にあらかじめ定めておき、何れかの書き込み要求がＣＰＵ１からアレイディスクコントローラ２に供給された時点で、その書き込み要求が指定する論理アドレスに対して、一つのパリティグループと、そのパリティグループに対して割り当てられた何れかのドライブの何れかの記憶領域を割り当てる。そのパリティグループに割り当てられた他の記憶領域に未だ何れのデータも保持されていないときには、そのパリティグループの他のデータは所定の値、例えば０とみなし、その書き込みデータからパリティデータを生成する。上記書き込みデータとこの生成されたパリティデータを互いに異なるドライブに書き込む。

【０１２８】次に来た第２の書き込み要求に対して同じパリティグループが割り当てられたとき、そのパリティグループに対して生成されたパリティデータ（旧パリティデータを読み出し、先にその第２の書き込み要求が指定する第２のデータとこの旧パリティから新たにパリティを生成し、この第２の書き込みデータをそのデータに割り当てられた記憶位置に書き込むとともに、生成されたパリティでもって、旧パリティを書き直す。その後も同様にしてそのパリティグループの全データに対する書き込みを実行する。このようにＲＡＩＤ５では、ＲＡＩＤ３と異なり、データの書き込みごとに、そのデータが属するパリティグループの旧パリティデータとその書き込みデータから新パリティデータを生成するが、このパリティの生成にそのグループの他のデータは使用しない。書き込み済みのデータを読み出すときには、そのデータを何れか一つのドライブから読み出すが、そのデータが属するパリティグループの他のデータは読み出さない。

【０１２９】（ドライブ障害があるときの装置動作）このときの動作はＲＡＩＤ３とＲＡＩＤ５との差に由来する違いを除いて実施例１と基本的には同じである。例えば、障害ドライブを有するボードがマザーボードに接続されている状態では、ＣＰＵ１からの、正常なドライブへのデータの書き込み要求を実行する場合には、上記ドライブ障害がない場合に説明したのと同様に、この書き込みデータを分割する必要がない。また、この障害ドライブを有するバードがマザーボードから分離されている間に、この分離されているボード上の障害ドライブに対し上位装置から書き込み要求が発行された場合、この書き込み要求に付随する書き込みデータは一時的にキャッシュメモリ４に保持しておく。

【０１３０】その障害ドライブが正常な交替ドライブと置換された後、そのボードがマザーボードと再度接続されたときに、交替ドライブが保持すべきデータを回復し、回復されたデータをその交替ドライブに書き込み、その後、キャッシュメモリに保持された書き込みデータを、その書き込みデータを分割することなくその交替ドライブに書き込む。障害ドライブを有するボード内の正常なドライブに関しても、このデータ回復を行なう必要がないことを除いて、同様である。

【０１３１】（実施例３の変形例）実施例１の各変形例がこの実施例３にも適用できる。ただし、ＲＡＩＤ３とＲＡＩＤ５との違いに由来する変更が必要である。

【０１３２】例えば、実施例１の変形例１では、障害ドライブを有するボードがマザーボードから分離されている状態で障害ドライブに対する書き込み要求がＣＰＵ１から供給された場合、書き込みデータを複数のサブデータに分割し、それからパリティデータを生成し、これらのサブデータの内、障害ドライブに書き込むべきデータ

以外のサブデータおよび生成された新パリティデータをそれぞれ異なるドライブに書き込んだ。実施例１のこの変形例１に対応する本実施例３の変形例では、書き込みデータを分割する必要はない。すなわち、障害ドライブを有するボードがマザーボードから分離されている間に上位装置から書き込みデータが供給されたとき、その書き込みデータが属するパリティグループの複数のデータの内、その障害があるドライブ以外のドライブに保持された複数のデータおよびそのパリティグループに属するパリティデータを読み出し、これらの読み出されたデータと新たにＣＰＵ１から供給された書き込みデータとから、新パリティを生成する。生成された新パリティデータでもって旧パリティを書き換えるだけでよい。

【０１３３】また、ＲＡＩＤ４のアレイコントローラにも本実施例は適用できる。

【０１３４】（実施例４）本実施例は本願発明を実施例２の装置の制御にＲＡＩＤ５を採用した実施例である。したがって、本実施例は、実施例２と同様に図１に示すものに予備ドライブボード５０Ａを付加したものと同じ装置構成により実現されるが、ＭＰ１７による制御が実施例２と異なる。したがって、本実施例の説明でも図１などの回路を使用して説明する。なお、説明は実施例２、３と異なる点を主にする。

【０１３５】（ドライブ障害がないときの装置動作）このときの動作はＲＡＩＤ３とＲＡＩＤ５との差に由来する違いを除いて実施例２と基本的には同じである。その違いは、実施例３と実施例１との違いに関して説明したとおりである。

【０１３６】（ドライブ障害があるときの装置動作）このときの動作はＲＡＩＤ３とＲＡＩＤ５との差に由来する違いを除いて実施例２と基本的には同じである。その違いは、実施例３と実施例１との違いに関して説明したとおりである。

【０１３７】（実施例４の変形例）実施例２の各変形例がこの実施例４の変形例としても使用できる。ただし、ＲＡＩＤ３とＲＡＩＤ５との違いに由来する変更が必要である。

【０１３８】例えば、障害ドライブへのデータの書き込みに関するＣＰＵ１からの書き込み要求を予備ドライブを使用して実行する場合には、上記ドライブ障害がない場合と同様に、この書き込みデータを分割する必要がない。すなわち、その書き込みデータが属するパリティグループの複数のデータの内、その障害があるドライブ以外のドライブに保持された複数のデータおよびそのパリティグループに属するパリティデータを読み出し、これらの読み出されたデータと新たにＣＰＵ１から供給された書き込みデータとから、その書き込みデータに対応する新パリティデータを生成し、この書き込みデータを予備ドライブに書き込み、この新パリティデータでもって旧パリティデータを書き換える。

【０１３９】また、ＲＡＩＤ４のアレイコントローラにも本実施例は適用できる。

【０１４０】（実施例５）以上の実施例と異なり、本実施例はドライブに障害がないがボードのコネクタに障害が発生した場合の実施例である。

【０１４１】何れかのドライブボードのコネクタがマザーボードとの接触不良などの不良が発生した場合にも、そのボード上の全ドライブはアクセスできなくなる。その場合、このボードを正常なものと交換するまで、そのボードをマザーボードから抜き取ることが必要になる。本実施例では、このような場合に、以上の実施例のいずれかを適用したものである。

【０１４２】すなわち、ユーザは何れかのボードのコネクタの接触異常を発見したとき、ＣＰＵ１を介してアレイコントローラ２に、そのボードを抜き取った通知するコマンドを入力する。その後、そのボードを抜き取る。この後の装置の動作は、上記実施例の何れかに記載のとおりである。

【０１４３】なお、何れかのドライブボードのコネクタの接触不良は次のような場合に検出可能である。そのドライブボード内のドライブをアクセスした結果、そのアクセスが不成功に終わると、アレイコントローラ２は、そのボードがマザーボードから分離されていない状態では、そのドライブに障害が発生したことを示す障害フラグをアドレステーブルにセットする。このようにして同じボードのドライブが次々と障害が発生したと判断された場合、一般には、それらのドライブの障害よりも別のところの障害の可能性が高い。その障害の原因の一つはそのボードのコネクタの接触不良である。ユーザによる調査の結果、そのコネクタの接触不良が判明したときには、ユーザは上記のコマンドを入力する。それに先立ち、すでに障害が発生したとされたドライブがあれば、それらのドライブに対する障害ビットをオフにするコマンドをユーザが入力する。アレイコントローラ２は、このコマンドの入力に応答して、そのコマンドで指定されるボード上の障害とされたドライブ上のドライブに対する障害ビットをリセットする。

【０１４４】

【発明の効果】以上説明したように、本発明によるディスクアレイ装置は、障害が発生したドライブを正常なドライブに交換する際、障害が発生したドライブが搭載されている基板をマザーボードから外して交換している間も使用し続けることができる。また、何れかのドライブボードのコネクタに障害が発生しても、そのボード上のドライブ内のデータを指定するデータ読み出し要求およびデータ書き込み要求を実行できる。

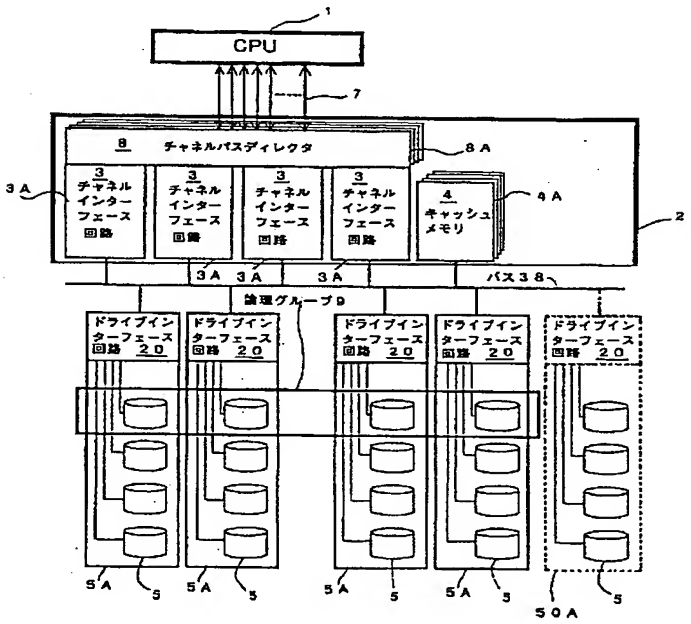
【図面の簡単な説明】

【図１】本発明に係るディスクアレイシステムの第１の実施例の概略的な装置構成を示す図

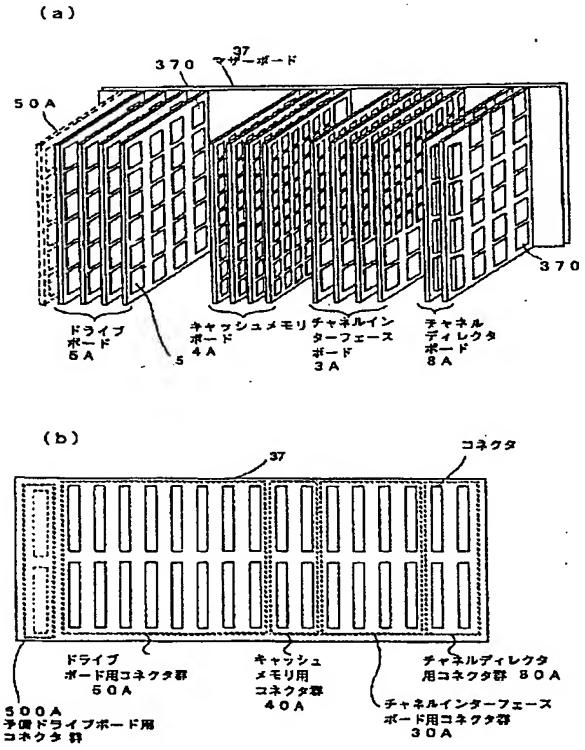
【図２】図１の装置のボードの配置およびマザーボード

のコネクタを示す図
【図３】図１の装置のより詳細な回路図
【図４】図３の回路に使用する論理グループテーブル（２１）の一例を示す図
【図５】図３の回路に使用するアドレステーブル（３１）の一例を示す図
【図６】第１の実施例によるデータの書き込みを説明する図
【図７】本発明に係るディスクアレイシステムの第２の実施例で使用する論理グループテーブルの一例を示す図
【図８】第２の実施例で使用するアドレステーブルの一例を示す図
【図９】本発明に係るディスクアレイシステムの第３の実施例で使用するアドレステーブルの一例を示す図
【符号の説明】
２…アレイコントローラ、３Ａ…チャネルインターフェースボード（ＣＩＢ）、４Ａ…キャッシュメモリボード、５Ａ…ドライブボード、７…チャネルパス、８…チャネルパスディレクタ、９…論理グループ、１０…インターフェースアダプタ、１１…チャネルパススイッチ、１２…データ線、１３…チャネルインターフェース（ＣＨ　ＩＦ）回路、１４…データ制御回路（ＤＣＣ）、１５…チャネル側キャッシュアダプタ（Ｃ　Ａｄｐ）、１６…制御信号線、１７…マイクロプロセッサ（ＭＰ）、１８…パリティ生成回路（ＰＧ）、１９…ドライブ側キャッシュアダプタ（Ｃ　Ａｄｐ）、２０…ドライブインターフェース回路（Ｄｒｉｖｅ　ＩＦ）、２１…論理グループテーブル、２２…論理アドレス（論理Ａｄｄｒ）、２３…論理グループＮｏ、２４…サブデータ名、２５…パリティ名、２７…ボードアドレス（ボードＡｄｄｒ）、２８…Ｄｒｉｖｅ　Ｎｏ、２９…障害フラグ２９、３０…ＳＣＳＩドライブアドレス、３１…アドレステーブル、３２…Ｄｒｉｖｅ内アドレス、１００…ボード抜去フラグ。

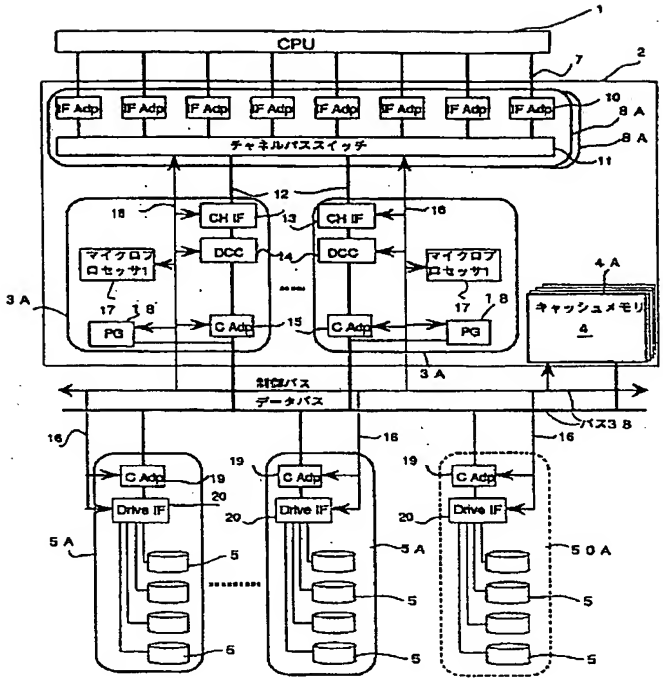
【図１】



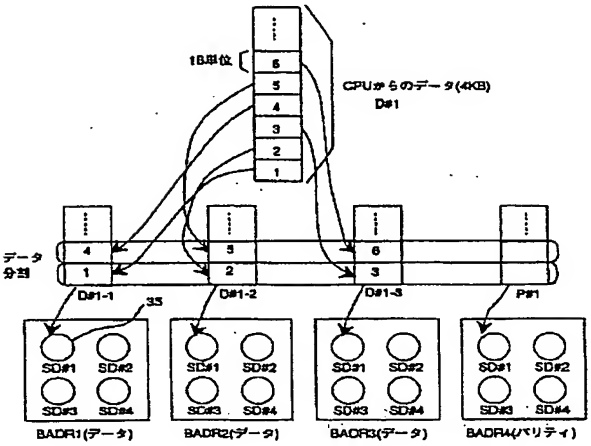
【図２】



【図３】


CPU
IF Adp
チャネルパススイッチ
CH IF
DCC
マイクロプロセッサ1
PG
C Adp
キャッシュメモリ
4
制御パス
データパス
Drive IF

【図６】

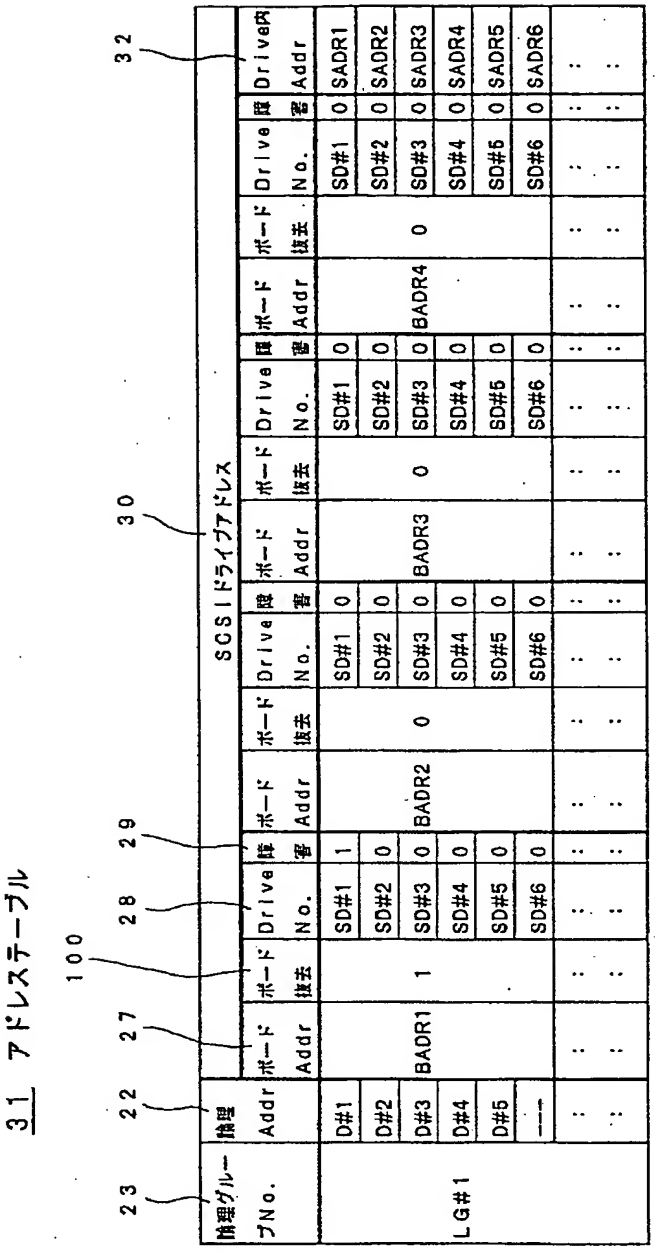

1Bブロック
CPUからのデータ(4KB)
D#1
データ分割
D#1-1
D#1-2
D#1-3
P#1
SD#1
SD#2
SD#3
SD#4
BADR1(データ)
BADR2(データ)
BADR3(データ)
BADR4(パリティ)

【図４】

論理グループテーブル

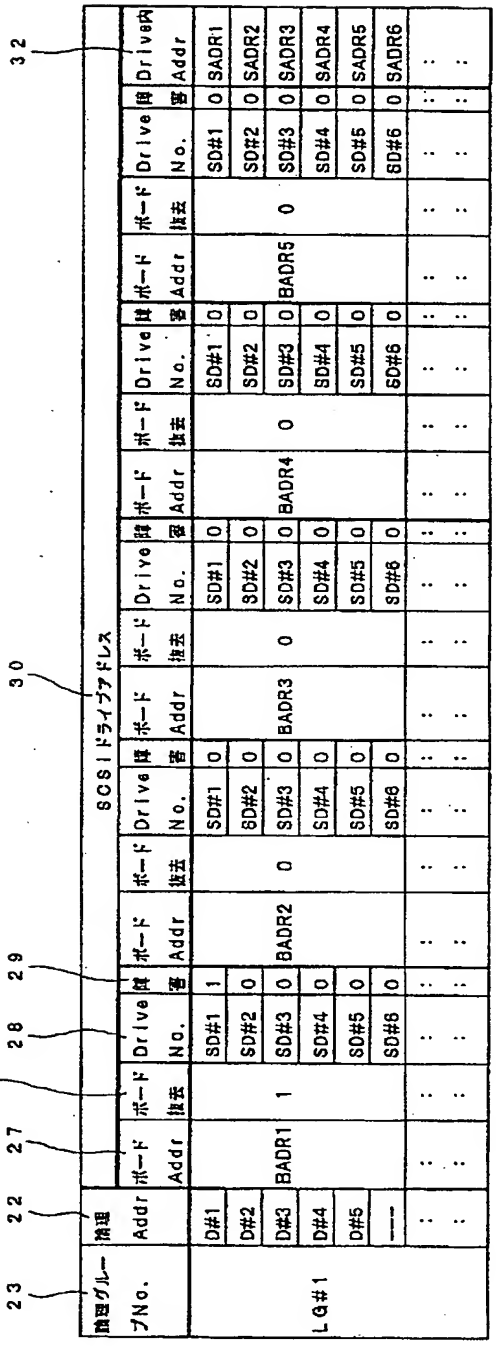
22 論理Addr	23 論理グループNo.	21 サブデータ名	24		25 パリティ名
D#1	LG#1	D#1-1	D#1-2	D#1-3	P#1
D#2		D#2-1	D#2-2	D#2-3	P#2
D#3		D#3-1	D#3-2	D#3-3	P#3
D#4		D#4-1	D#4-2	D#4-3	P#4
D#5		D#5-1	D#5-2	D#5-3	P#5
———		———	———	———	———
⋮		⋮	⋮	⋮	⋮
D#11	LG#2	D#11-1	D#11-2	D#11-3	P#11
D#12		D#12-1	D#12-2	D#12-3	P#12
D#13		D#13-1	D#13-2	D#13-3	P#13
D#14		D#14-1	D#14-2	D#14-3	P#14
D#15		D#15-1	D#15-2	D#15-3	P#15
⋮		⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮

【図５】


31 アドレステーブル
100
23
22
27
28
29
30
32
論理グループNo.
論理Addr
ボードAddr
ボード抜去
Drive No.
障害
ボードAddr
ボード抜去
Drive No.
障害
ボードAddr
ボード抜去
Drive No.
障害
ボードAddr
ボード抜去
Drive No.
障害
Drive内Addr
SCSIドライブアドレス
LG#1
D#1
D#2
D#3
D#4
D#5

BADR1
1
SD#1
SD#2
SD#3
SD#4
SD#5
SD#6
0
BADR2
BADR3
BADR4
SADR1
SADR2
SADR3
SADR4
SADR5
SADR6

【図８】


31 アドレステーブル
100
23
22
27
28
29
30
32
論理グループNo.
論理Addr
ボードAddr
ボード抜去
Drive No.
障害
ボードAddr
ボード抜去
Drive No.
障害
ボードAddr
ボード抜去
Drive No.
障害
ボードAddr
ボード抜去
Drive No.
障害
Drive内Addr
SCSIドライブアドレス
LG#1
D#1
D#2
D#3
D#4
D#5

BADR1
1
SD#1
SD#2
SD#3
SD#4
SD#5
SD#6
0
BADR2
BADR3
BADR4
BADR5
SADR1
SADR2
SADR3
SADR4
SADR5
SADR6

【図７】

論理グループテーブル

22 論理Addr	23 論理グループNo.	21 サブデータ名	24		25 パリティ名	26
D#1	LG#1	D#1-1	D#1-2	D#1-3	P#1	Spare1
D#2		D#2-1	D#2-2	D#2-3	P#2	Spare2
D#3		D#3-1	D#3-2	D#3-3	P#3	Spare3
D#4		D#4-1	D#4-2	D#4-3	P#4	Spare4
D#5		D#5-1	D#5-2	D#5-3	P#5	Spare5
——		——	——	——	——	——
⋮		⋮	⋮	⋮	⋮	⋮
D#11	LG#2	D#11-1	D#11-2	D#11-3	P#11	Spare11
D#12		D#12-1	D#12-2	D#12-3	P#12	Spare12
D#13		D#13-1	D#13-2	D#13-3	P#13	Spare13
D#14		D#14-1	D#14-2	D#14-3	P#14	Spare14
D#15		D#15-1	D#15-2	D#15-3	P#15	Spare15
⋮		⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮

【図９】

31　アドレステーブル

23	22	27	30	28	29	27	100	28	29	32
論理グループ No.	論理 Addr	SCSIドライブアドレス								
		データ				パリティ				
		ボード Addr	ボード 複写	Drive No.	障害	ボード Addr	ボード 複写	Drive No.	障害	Drive内 Addr
LG#1	D#1	BADR1	0	SD#1	0	BADR4	0	SD#1	0	SADR1
	D#2	BADR2	0							
	D#3	BADR3	0							
	D#4	BADR2	0	SD#1	0	BADR1	0	SD#1	0	SADR2
	D#5	BADR3	0							
	D#6	BADR4	0							
	D#7	BADR3	0	SD#1	0	BADR2	0	SD#1	0	SADR3
	D#8	BADR4	0							
	D#9	BADR1	0							
	D#10	BADR4	0	SD#1	0	BADR3	0	SD#1	0	SADR4
	D#11	BADR1	0							
	D#12	BADR2	0							
	D#13	BADR1	0	SD#2	0	BADR4	0	SD#2	0	SADR1
	D#14	BADR2	0							
	D#15	BADR3	0							
	D#16	BADR2	0	SD#2	0	BADR1	0	SD#2	0	SADR2
	D#17	BADR3	0							
	D#18	BADR4	0							
	D#19	BADR3	0	SD#2	0	BADR2	0	SD#2	0	SADR3
	D#20	BADR4	0							
	D#21	BADR1	0							
	D#22	BADR4	0	SD#2	0	BADR3	0	SD#2	0	SADR4
	D#23	BADR1	0							
	D#24	BADR2	0							
	:	:	:	:	:	:	:	:	:	:
	:	:	:	:	:	:	:	:	:	: